

UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS

École doctorale de STIC

THÈSE

pour obtenir le titre de

Docteur en Sciences

de l'UNIVERSITÉ de Nice-Sophia Antipolis

Mention : Automatique, Traitement du Signal et des Images

présentée et soutenue par

Muriel GASTAUD

MODÈLES DE CONTOURS ACTIFS POUR LA SEGMENTATION D'IMAGES ET DE VIDÉOS

Thèse dirigée par Pr. Michel BARLAUD et Pr. Gilles AUBERT

soutenue le 6 décembre 2005

Jury :

Gilles AUBERT	Professeur des Universités (Nice-Sophia Antipolis)	Président
Patrick BOUTHEMY	Directeur de recherche INRIA à l'IRISA (Rennes)	Rapporteur
Henri MAITRE	Professeur des Universités à l'ENST (Paris)	Rapporteur
Fernand MEYER	Directeur de recherche à l'ENSMP (Paris)	Rapporteur
Nikos PARAGIOS	Professeur à l'ECP (Paris)	Examineur
Michel BARLAUD	Professeur des Universités (Nice-Sophia Antipolis)	Directeur de thèse

Remerciements

En premier lieu, je tiens à remercier l'ensemble des membres du jury pour m'avoir fait l'honneur de leur présence et pour l'intérêt qu'ils ont porté à mes travaux.

Je remercie monsieur Gilles Aubert pour avoir accepté de présider ce jury, messieurs Patrick Bouthemy, Henri Maitre et Fernand Meyer, pour leur relecture attentive du manuscrit. Leurs commentaires m'ont été très précieux et ont largement contribué à l'amélioration de ce document. Je remercie monsieur Nikos Paragios pour avoir accepté de faire partie du jury.

Je remercie messieurs Pierre Bernhard et Jean-Marc Fédou, successivement directeurs du laboratoire I3S, pour m'avoir accueillie au sein de l'équipe CReATIVe.

Je remercie monsieur Michel Barlaud, responsable de l'équipe CReATIVe, pour avoir accepté de diriger ma thèse. Je le remercie pour son enthousiasme et pour l'implication dont il fait preuve à l'égard de nos travaux et de notre devenir. Je le remercie pour l'attention qu'il porte à chacun de ses thésards. Je remercie monsieur Gilles Aubert pour avoir co-dirigé ma thèse et pour m'avoir fait bénéficier de son expérience, de sa rigueur scientifique, et pour son apport mathématique, élément essentiel à ce manuscrit.

Je tiens à remercier monsieur Daniel Gaffé, mon tuteur de monitorat, monsieur Michel Barlaud, monsieur Vincent Granet, et madame Céline Theys Ferrari pour avoir guidé mes premiers pas dans le monde enseignant. La diversité des domaines et des niveaux d'enseignement qu'ils m'ont permis d'effectuer ont rendu cette expérience très enrichissante.

J'y associe Stéphanie, Frédéric, Thomas, Sylvain et Stéphane pour m'avoir accompagnée dans cette aventure, ainsi que l'ensemble de mes élèves.

Je souhaite remercier Micheline Hagnéré, notre assistante de projet, pour son

efficacité, sa gentillesse et son sourire. J’y associe l’ensemble du personnel administratif pour leur disponibilité et leur bonne humeur.

Je remercie Guy Tessier pour sa gentillesse et son franc-parler.

Je remercie Eric pour avoir redonné du souffle à cette seconde partie de thèse, pour ses relectures minutieuses et ses corrections, ainsi que la diplomatie avec laquelle il sait rendre un écrit rouge (ou vert) de «suggestions». J’y associe l’ensemble des personnes qui, par leurs relectures et leurs remarques constructives, ont contribué à améliorer ce manuscrit, et plus particulièrement Frédéric, toujours présent quand on a besoin de lui, Marie, mon coach officiel *es* perspectives, mon Thomas, traducteur d’expressions franco-mentonnaises, Ariane, contrainte de lire et de relire la contrainte géométrique, et Vincent détecteur d’erreurs dans la détection de mouvement.

Je remercie de tout cœur l’ensemble des personnes qui ont contribué à l’esprit CReATiVe :

Annabelle, Arnaud, Christophe, Frédéric et Marie, Frédéric et Sabine, Lionel, Manuela et Simão (et Patricia), Marco, Olivier, Stéphane, Stéphanie, Valéry ;

ou qui y contribuent encore :

Akram, Ariane, Blaise, Eric et Asun, Leonardo, Marie, Marie-Andrée, Sylvain, Thomas, Vincent, Yasmine.

Merci pour les pauses-café, les randos, le roller, les journées au ski, ..., les fêtes d’anniversaire et toutes les bonnes occasions de faire un repas entre amis, votre écoute attentive de mes coups de blues et de mes coups de gueule, et votre indulgence envers mon humour. Vous êtes la seule raison pour laquelle je continue de manger au resto U. Votre amitié m’est précieuse. J’y associe l’ensemble des stagiaires qui ont renforcé l’équipe, et les thésards du deuxième étage, qu’ils me pardonnent de ne pas les nommer tous.

Petite pensée émue pour les deux autres membres de la famille 126 dispersée aux quatre coins de l’hexagone.

Je remercie Thomas, en ce qui concerne cette thèse, pour son aide, sa patience, ses encouragements et son soutien quotidien. Je joins à ces remerciements sa famille : l’affection et l’attention qu’ils m’ont accordées, en particulier en période de soutenance, m’ont beaucoup touchée.

Enfin, je remercie mes parents, ma sœur et ma famille, pour leur amour et leur soutien inconditionnels. Je leur dédie cette thèse. Merci d’avoir toujours cru en moi. En retour, je promets à «Junior» que sa tata le gâtera, et lui souhaite, par avance, la bienvenue...

Table des matières

1	Introduction	1
1.1	Introduction à la segmentation d'images	1
1.2	Structure du manuscrit	4
I	Introduction aux contours actifs	9
2	État de l'art des contours actifs	11
2.1	<i>Snakes</i>	12
2.2	Contours actifs géométriques	13
2.3	Contours actifs géodésiques	14
2.4	Descripteurs de région	16
3	Approche variationnelle et gradients de forme	19
3.1	Descripteurs de contour et de région	21
3.2	Outils de dérivation	22
3.2.1	Introduction du schéma dynamique dans le critère . . .	22
3.2.2	Définition d'une dérivée de domaine et théorème d'équivalence	24
3.3	Dérivation de termes de contour et de région	25
3.3.1	Dérivation d'un terme de contour	25
3.3.2	Dérivation d'un terme de région dont le descripteur ne dépend pas de la région	26
3.4	Dérivation d'un terme de région à double dépendance	28
3.4.1	Exemples de dérivation d'un terme de région à double dépendance	28
3.4.2	Cas général : théorème	35
4	Mise en œuvre des contours actifs	37
4.1	Contours actifs paramétriques	37

4.2	Méthode des ensembles de niveaux	39
II Segmentation avec a priori géométrique sans contrainte paramétrique		43
5	Contours actifs et a priori géométrique	45
5.1	Introduction et état de l'art	45
5.2	Définition de la contrainte géométrique non paramétrique . . .	47
5.2.1	Distance à un contour : définition et propriétés directes	47
5.2.2	Définition du critère d'a priori géométrique	48
5.3	Segmentation par contours actifs : une approche variationnelle	50
5.3.1	Dérivation du critère d'a priori géométrique	50
5.3.2	Démonstration par le calcul des variations	51
5.3.3	Démonstration par les gradients de forme	55
5.3.4	Équation d'évolution du contour actif	56
6	Applications	59
6.1	Déformation de contour	59
6.1.1	Choix de la fonction de pondération de la distance . . .	60
6.1.2	Mise en œuvre	61
6.1.3	Résultats expérimentaux	62
6.2	Segmentation interactive d'images	66
6.2.1	Critère de compétition de régions	66
6.2.2	Critère combinant a priori géométrique et descripteur de région	68
6.2.3	Résultats expérimentaux	70
6.3	Segmentation de vidéos	72
6.3.1	Estimation du contour de référence	72
6.3.2	Critère de segmentation de vidéos et équation d'évolu- tion du contour	74
6.3.3	Résultats expérimentaux	75
III Segmentation et estimation conjointes du mouve- ment		79
7	État de l'art	81
7.1	Détection de mouvement	82
7.2	Segmentation et suivi d'objets en mouvement	85
7.2.1	Segmentation et estimation séquentielle du mouvement	85

7.2.2	Segmentation et estimation conjointe du mouvement	86
8	Segmentation et estimation conjointes entre deux images	89
8.1	Définition du critère conjoint entre deux images	89
8.2	Segmentation et suivi d'objets en mouvement	92
8.2.1	Équation d'évolution pour la segmentation d'objets en mouvement	96
8.2.2	Équation d'évolution pour le suivi d'objets en mouvement	96
8.3	Estimation du mouvement entre deux images	97
8.3.1	Modèle de translation uniforme	99
8.3.2	Modèle de translation et d'homothétie centrée	100
9	Segmentation et estimation conjointes sur un groupe d'images	103
9.1	Définition du critère conjoint sur un groupe d'images	103
9.2	Segmentation et suivi d'objets en mouvement	105
9.2.1	Équation d'évolution pour la segmentation d'objets en mouvement	107
9.2.2	Équation d'évolution pour le suivi d'objets en mouvement	107
9.3	Estimation du mouvement	107
9.3.1	Modèle de translation uniforme	108
9.3.2	Modèle de translation uniforme et d'homothétie centrée	109
10	Résultats expérimentaux	111
10.1	Segmentation/estimation conjointes du mouvement	111
10.1.1	Validation de l'estimation du mouvement	111
10.1.2	Validation de la segmentation conjointe	115
10.2	Suivi d'objets en mouvement	116
10.2.1	Directions d'évolution du contour actif	116
10.2.2	Résultats en suivi d'objets en mouvement	118
IV	Conclusions et perspectives	121
V	Annexe : Détection d'objets en mouvement	129
A	Estimation du fond de la séquence : la mosaïque	135
A.1	Estimation du mouvement de la caméra	136
A.1.1	Estimation du mouvement par mise en correspondance	137

A.1.2	Estimation robuste du modèle de mouvement de la ca- méra	138
A.2	Compensation du mouvement de la caméra	142
A.3	Segmentation grossière du fond des images	143
A.4	Mise à jour de la mosaïque	145
B	Segmentation finale : détection d’objets en mouvement avec fond estimé	149
B.1	Extraction du fond de chaque image à partir de la mosaïque .	149
B.2	Détection d’objet en mouvement	149
C	Résultats expérimentaux	153

Table des figures

1.1	Segmentation : image originale.	2
1.2	Extraction des contours d'une image à partir de ses gradients.	2
1.3	Segmentation en régions à partir des contours morphologiques par <i>watershed</i> [5, 84]	3
1.4	Segmentation en régions de couleur homogène.	3
1.5	Segmentation par contours actifs : du contour initial (à gauche) au contour à convergence (à droite).	4
2.1	Principe des contours actifs : le contour actif Γ évolue vers l'ob- jet d'intérêt sous l'action d'une force dirigée suivant la normale \mathbf{N} au contour.	11
2.2	Évolution du contour actif suivant l'équation $\frac{\partial \Gamma}{\partial \tau} = -\nabla g(\nabla I(\Gamma)) \cdot \mathbf{N}\mathbf{N}$ près des frontières de l'objet.	16
3.1	Approche variationnelle avec calcul de l'EDP.	20
3.2	Étapes de dérivation d'un terme de région dont le descripteur est une fonction de la moyenne de la région.	29
3.3	Étapes de dérivation d'un terme de région dont le descripteur est une fonction de la variance de la région.	33
4.1	Mise en œuvre paramétrique : construction d'une B-spline cu- bique.	38
4.2	Méthode des ensembles de niveaux : gestion de la topologie. La première ligne d'images représente la carte des distances. Chaque carte est coupée en son niveau 0. Les contours corres- pondants constituent la deuxième ligne d'images.	40
5.1	Squelette de Γ^{ref} : pour une forme concave, le squelette a une composante intérieure et au moins une composante extérieure	48

5.2	Illustrations sur une image réelle : de gauche à droite, un contour de référence (en rouge) détournant un objet, la carte de distance signée associée, et le squelette intérieur de l'objet.	48
5.3	Fonction distance signée à un contour de référence	49
5.4	Normales et tangentes aux contours actifs et de référence. . . .	54
6.1	Fonction de pondération de la distance.	61
6.2	Déformation d'une ellipse à une autre	63
6.3	Déformation d'un contour d'un visage à un autre : définition des contours initial et de référence	64
6.4	Déformation d'un contour d'un visage à un autre : évolution du contour actif	65
6.5	Comparaison entre un algorithme de compétition de régions et l'algorithme de segmentation interactive incluant l'a priori géométrique.	71
6.6	Estimation du contour de référence pour le tracking	73
6.7	Segmentation de visage utilisant l'a priori géométrique sur 50 images de la vidéo Erik.	77
8.1	Modèle de mouvement entre deux images	91
8.2	Différentes directions d'évolution	97
9.1	Modèle de mouvement sur un groupe d'images.	104
10.1	Estimation du mouvement entre les images 80 et 79 de la séquence Foreman : à gauche, l'estimation par l'équation (10.1), à droite, l'estimation par <i>block-matching</i>	112
10.2	Histogrammes de l'erreur de prédiction pour l'image 80 : en rouge notre méthode, en bleu le <i>block-matching</i>	113
10.3	A gauche : l'image originale ; A droite : les vecteurs de déplacement obtenus en estimant le mouvement par blocs 15 * 15 sur 3 images.	114
10.4	Estimation du mouvement par région sur 3 images de la séquence Foreman.	115
10.5	Estimation et segmentation conjointes du mouvement de 2 images successives de la séquence Coastguard sur un groupe de 3 images. Sur la première ligne : le contour initial et la segmentation de l'image 250 ; Sur la deuxième ligne : <i>Idem</i> pour l'image 251.	116
10.6	Suivi d'objet en mouvement sur la séquence synthétique créée à partir de Foreman.	118

10.7	Suivi d'objets en mouvement sur 15 images de la séquence <i>flower and garden</i> avec l'équation d'évolution suivant la direction du mouvement.	119
8	Étapes-clés de l'algorithme de détection de mouvement à partir d'une mosaïque.	133
A.1	Estimation du mouvement dite <i>Backward</i>	137
A.2	L'estimateur de Geman et Mc Clure	140
A.3	Compensation <i>backward</i> du mouvement de la caméra entre les images I_n et I_{n-1}	142
A.4	Contours volontairement sur-segmentés des images 70, 75, 80 et 85.	145
A.5	Construction de la mosaïque à partir des fonds grossiers. . . .	146
B.1	Extraction des fonds de la séquence à partir de la mosaïque. .	150
B.2	Présentation du descripteur basé sur la différence entre l'image et son fond.	151
C.1	Mosaïque estimée à partir de 20 images de la séquence <i>Stefan</i> (images 70 à 89)	154
C.2	Fonds et objets des images 70,75,77, 80 et 87	155
C.3	Contours des images 70,72,74, 76, 78, 80, 82, 84, 86 et 88 . . .	156

Notations

Notations principales

- Les vecteurs sont représentés par des caractères gras : par exemple le point \mathbf{x}
- Les matrices sont représentées par des majuscules : par exemple la matrice M .
- Les images sont représentées par des fonctions continues $I : \Omega \rightarrow \mathcal{R}^m$, avec l'intensité lumineuse associée à chaque pixel de l'image, avec Ω un ensemble de \mathcal{R}^n représentant le support de l'image.
- Pour des images couleur, $m = 3$, l'intensité est représentée par une fonction $I : \Omega \rightarrow \mathcal{R}^m$ où $I = [I^1, I^2, I^3]^T$. I^1 représente la luminance et I^2 et I^3 les chrominance de l'image dans l'espace de couleurs YUV . Pour des images à niveaux de gris, $m = 1$, on notera alors l'intensité de l'image par I .
- Pour les séquences de N images, chaque image est représentée par une fonction $I_n : \Omega_n \rightarrow \mathcal{R}^3$ avec $n \in [0, N - 1]$.
- Soit $\Gamma(p) : [a, b] \rightarrow \mathcal{R}^2$ une famille de courbes fermées paramétrées par p . Le contour est orienté dans le sens inverse des aiguilles d'une montre. Une des paramétrisations possible du contour est l'abscisse curviligne, notée s définie de telle manière que $|\Gamma'(s)| = 1$. Dans ce cas, le contour $\Gamma(s)$ est défini sur $[0, L]$ où L est la longueur de la courbe.
- \mathbf{T} est la tangente unitaire au contour Γ .
- \mathbf{N} est le vecteur normal intérieur au contour Γ .
- κ est la courbure euclidienne du contour Γ .
- Ω^{in} et Ω^{out} sont respectivement la région intérieure au contour, soit l'objet, et la région extérieure au contour, soit le fond de l'image.
- La dérivée classique d'une fonction $f(r)$ est notée $f' : f'(r) = \frac{df(r)}{dr}$. Elle peut également être notée df .

- La dérivée au deuxième ordre est notée : $f''(r)$ ou d^2f .
- La dérivée partielle d'une fonction à plusieurs variables, par exemple $f(a, b)$, est notée $\frac{\partial f}{\partial a}(a, b) = \lim_{\tau \rightarrow 0} \frac{f(a+\tau, b) - f(a, b)}{\tau}$.
 - La dérivée partielle au deuxième ordre est notée $\frac{\partial^2 f}{\partial a^2}$.
 - $\nabla = [\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}]^T$ est l'opérateur gradient.
 - La dérivée d'une fonction $J(x)$ dans la direction \mathbf{V} est notée $dJ(x, \mathbf{V})$ (voir la définition 3.1 page 24).
 - La dérivée de domaine d'une fonction $k(x, \Omega)$ dans la direction \mathbf{V} est notée $d_\Omega k(x, \Omega, \mathbf{V})$ (voir la définition 3.3 page 25).
 - La dérivée matérielle d'une fonction $k(x, \Omega)$ dans la direction \mathbf{V} est notée $\dot{k}(x, \Omega, \mathbf{V})$ (voir la définition 3.2 page 24).
 - Les noms de séquences vidéos sont écrits sans sérif : par exemple **Foreman**.
 - Le produit scalaire de deux vecteurs \mathbf{x} et \mathbf{y} est noté $\mathbf{x} \cdot \mathbf{y}$, ou $\langle \mathbf{x}, \mathbf{y} \rangle$.
 - Les anglicismes sont écrits en italique : par exemple *shape warping*.

Acronymes et abréviations

- COST211 : *European COST211 Group - Research on Redundancy Reduction Techniques and Content Analysis for Multimedia Services* - Groupe de recherche européen sur les techniques de réduction des redondances temporelles et l'analyse du contenu pour les services multimédia.
- CReATIVE : Compression, Reconstruction, Adaptées au Traitement d'Images et à la Vidéo.
- DFD : *Displaced frame difference* - Différence pixel à pixel entre deux images, dont l'une est compensée en mouvement.
- EDP : Équations aux Dérivées Partielles.
- FD : *frame difference* - Différence pixel à pixel entre deux images.
- ENSMP : École Nationale Supérieure des Mines de Paris.
- ENST : École Nationale Supérieure des Télécommunications.
- H261, H263, H26L : normes de codage vidéo adopté par l'ITU-T.
- I3S : Informatique Signaux et Systèmes de Sophia Antipolis.
- INRIA : Institut National de Recherche en Informatique et en Automatique.
- IRISA : Institut de Recherche en Informatique et Systèmes aléatoires.
- ITU-T : *International Telecommunication Union, Telecommunication normalization sector* - Union internationale des télécommunications, organisme de normalisation qui développe des standards dans le do-

- maine des télécommunications.
- MPEG : *Moving Picture Expert Group* - Groupe d'experts chargés de mettre au point des formats de compression vidéo.
- MPEG1, MPEG2, MPEG4, MPEG7 : normes de codage vidéo adoptées par le groupe MPEG.
- OSIAM : Projet du RNRT, Outils pour la Segmentation d'Images Animés pour MPEG4/7.
- pixel : contraction de *picture element* - plus petite unité indivisible d'une image numérisée.
- Pr. : Professeur.
- RNRT : Réseau National de la Recherche en Télécommunications.
- SSD : *Sum of Squared Differences* - critère de comparaison entre deux régions fondé sur la somme, sur la région, des différences au carré.
- STIC : Sciences et Technologies de l'Information et de la Communication.
- UFR : Unité de Formation et de Recherche.

Anglicismes

Certains anglicismes lexicaux apparaissent dans ce document. Nous les traduisons dans le contexte de nos travaux.

- *block-matching* : méthode de mise en correspondance de bloc.
- *backward* : en arrière - désigne les méthodes qui utilisent l'image courante et l'image précédente.
- *forward* : en avant - désigne les méthodes qui utilisent l'image et courante et l'image suivante.
- *in, out* : à l'intérieur, à l'extérieur - désigne la région intérieure ou extérieure au contour.
- *outlier* : Pour un ensemble de points donnés, un *outlier* est un point qui se situe loin des autres. Dans un contexte statistique, les données sont réparties suivant une fonction de distribution. Un point qui se démarque de cette répartition est un *outlier*.
- *shape warping* : déformation de courbe.
- *snake* : (littéralement : serpent) désigne les premiers contours actifs introduits par Kass *et al.* [52].
- *tracking* : suivi de cibles, la cible peut être une région homogène ou un objet en mouvement.
- *watershed* : ligne de partage des eaux.

Chapitre 1

Introduction

1.1 Introduction à la segmentation d'images

Une image est constituée d'un ensemble de pixels.

Segmenter une image en régions, c'est rassembler ces pixels en ensembles ayant des propriétés communes, comme une couleur ou une texture similaire, ou un mouvement cohérent.

Segmenter une image en objets, c'est rassembler des régions qui ont un sens sémantique commun, même si leurs propriétés sont différentes. Il s'agit par exemple de distinguer un personnage du contexte dans lequel il évolue, alors que cette personne a des vêtements de couleurs différentes ou ne bouge que certaines parties de son corps.

L'importance de la segmentation croît avec celle de l'image dans notre société. En imagerie médicale, la segmentation peut aider le médecin dans son diagnostic. En compression vidéo, elle permet de traiter différemment une zone d'intérêt, qui bénéficiera d'une plus grande précision, du reste de l'image qui pourra être plus fortement compressé. En indexation, la segmentation sert à extraire un objet que l'on souhaiterait retrouver dans d'autres images. En post-production cinématographique, un personnage segmenté peut être remplacé dans un autre décor, ce qui est le cas notamment pour les présentateurs de la météo. En vidéo-surveillance, la détection d'objets en mouvement peut révéler la présence d'intrus, ou évaluer la fluidité d'un trafic routier. Les applications sont nombreuses et la liste des exemples cités est loin d'être exhaustive.

Mais revenons au principe de la segmentation. Pour segmenter une région, une méthode consiste à détecter son contour. Le passage d'une région à une autre se caractérise par une différence d'intensité lumineuse. Plus la différence est importante, plus le gradient de l'image sera fort, et plus le



FIG. 1.1 – Segmentation : image originale.

contour sera marqué. Prenons pour exemple l'image de la figure 1.1 et essayons de segmenter la coccinelle.

La figure 1.2 nous présente deux extractions possibles des gradients. Ces gradients définissent des contours qui n'ont pas tous le même intérêt pour nous. Les contours de la feuille sont bien marqués, alors que le contour de la

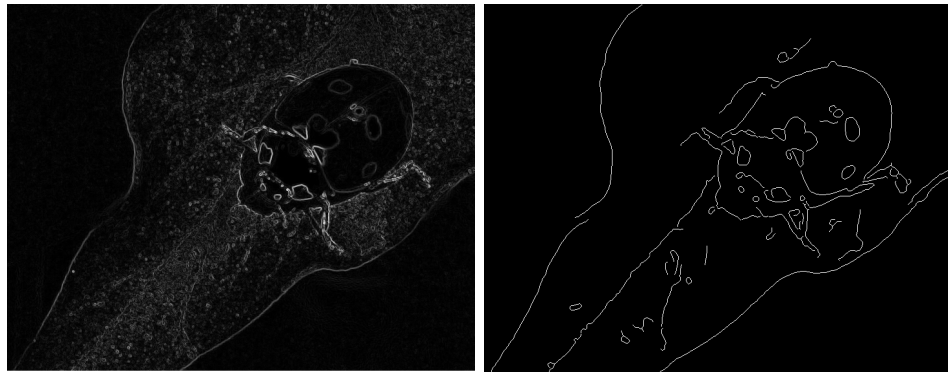


FIG. 1.2 – Extraction des contours d'une image à partir de ses gradients.

coccinelle n'a pas une intensité constante.

Si l'on peut déterminer l'intensité à laquelle les gradients sont détectés, on ne peut cependant pas déterminer leur nombre, s'assurer de leur connexité, ou leur donner un sens quant à leur appartenance à l'objet. Autrement dit, tous les gradients détectés n'appartiennent pas au contour de la coccinelle, et l'on ne saurait les faire disparaître sans perdre de la même manière une partie des contours de la coccinelle. De plus, ces contours ne sont pas

forcément fermés et ne définissent donc pas des régions séparées.

Certains algorithmes permettent de définir des régions à partir des gradients de l'image. Là encore, on peut contrôler le nombre de régions qui découpent



FIG. 1.3 – Segmentation en régions à partir des contours morphologiques par *watershed* [5, 84]

l'image, mais on ne peut pas savoir si ces régions appartiennent à l'objet d'intérêt ou non. Sur la figure 1.3, la coccinelle est bien détournée, mais certaines régions de la coccinelle (les régions blanches) sont considérées comme n'appartenant pas à la coccinelle. D'autres régions sont segmentées, mais elles ne nous intéressent pas.

Une autre approche consiste à définir directement les régions que l'on souhaite segmenter, à partir d'une caractéristique commune. Par exemple,



FIG. 1.4 – Segmentation en régions de couleur homogène.

on peut choisir de segmenter l'image en régions de même couleur, ou plus précisément de couleur homogène (voir figure 1.4). On rencontre le même problème que précédemment, à savoir que notre coccinelle est constituée de plusieurs régions de différentes couleurs et qu'on ne l'a toujours pas distinguée du reste de l'image.

La dernière approche que nous verrons consiste à définir un contour fermé et à le faire évoluer vers l'objet d'intérêt (voir figure 1.5). La tâche est difficile

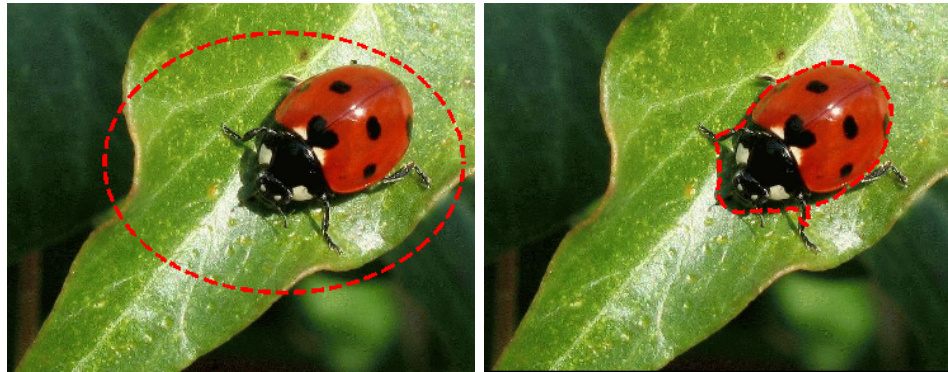


FIG. 1.5 – Segmentation par contours actifs : du contour initial (à gauche) au contour à convergence (à droite).

et nécessite d'avoir une idée de l'objet ou de la région que l'on souhaite segmenter. Pour définir cet objet, il faut choisir un critère qui détermine si un pixel fait partie de l'objet ou non. Ce critère doit contenir la description des propriétés de l'objet. Le contour actif peut par exemple évoluer vers des zones de fort gradient, tout en essayant d'entourer une région la plus homogène possible au niveau de la couleur.

De nombreuses propriétés peuvent être prises en compte dans le critère, et le contour actif évolue en effectuant un compromis entre elles. En cas d'objets multiples, il peut également se séparer ; au contraire, deux contours peuvent fusionner.

1.2 Structure du manuscrit

Nous présenterons les principes généraux des contours actifs, de la définition de leur critère à leur implémentation, en passant par le calcul de l'équation d'évolution.

Nous étudierons, dans ce cadre de travail, deux critères particuliers. Le premier critère définira un terme d'a priori géométrique minimisant la distance à un contour de référence. Seul ou couplé à un terme statistique, ce terme servira dans le cadre d'algorithmes de déformation de contours ou *shape warping*, de segmentation, et de suivi de régions homogènes ou *tracking*. Le deuxième exemple définira un terme de segmentation joint à l'estimation d'un modèle de mouvement, et sera appliqué à la segmentation et au suivi d'objets en mouvement.

Ce manuscrit est organisé en trois parties : la première partie est consacrée à l'étude théorique des contours actifs, et les deux autres à l'étude des critères particuliers.

La première partie est théorique et propose une **introduction aux contours actifs**. Elle est composée des chapitres suivants :

Chapitre 2 : État de l'art des contours actifs. Ce chapitre présente des méthodes classiques de contours actifs, que leur approche du problème soit variationnelle ou non. Nous présentons les *snakes*, travaux pionniers de Kass *et al.*, puis les contours actifs géométriques, et les contours actifs géodésiques. Si ces approches mettent en oeuvre des caractéristiques globales du contour, comme sa longueur ou sa rigidité, elles n'utilisent que des propriétés locales de l'image, comme son intensité ou son gradient en un point. Les dernières méthodes de contours actifs présentées dans ce chapitre s'attachent à caractériser les régions délimitées par le contour, et utilisent des propriétés globales à ces régions, comme leur moyenne, leur variance, leur entropie, leur histogramme ou leur mouvement.

Chapitre 3 : Approche variationnelle : dérivation par les gradients de forme. Ce chapitre présente le cadre de travail de notre approche et les principaux outils mis en oeuvre. Notre approche est variationnelle, c'est-à-dire que le problème de segmentation est formulé comme la minimisation d'un critère. La dérivation du critère permet de déduire l'équation d'évolution du contour actif. Dans ce chapitre, nous définirons deux types de termes pouvant constituer un critère de segmentation : des termes de contour et des termes de région, parmi lesquels des termes de régions dits à *double dépendance*. Puis, nous présenterons les outils de dérivation mis en oeuvre pour le calcul de l'équation d'évolution : les gradients de forme, outils issus de l'optimisation de domaine. Enfin, nous dériverons chacun des termes présentés, en portant une attention particulière à la dérivation des termes de région à double dépendance.

Chapitre 4 : Implémentation. Ce chapitre traite de l'implémentation de l'équation d'évolution du contour actif. Deux approches sont brièvement présentées : une approche paramétrique avec l'implémentation par B-Spline, et une approche implicite avec la méthode des ensembles de niveaux.

La deuxième partie du document est consacrée à la définition et l'étude d'un critère de **segmentation avec a priori géométrique sans contrainte paramétrique**. Plusieurs applications de ce critère sont proposées. La deuxième partie est composée des chapitres suivants :

Chapitre 5 : Contours actifs et a priori géométrique non paramétrique. Dans ce chapitre, nous définissons un critère d'a priori géométrique minimisant la distance euclidienne du contour actif à un contour de référence. Nous ne faisons pas l'hypothèse d'une transformation paramétrique entre le contour actif et le contour de référence, et notre critère ne dépend pas d'une paramétrisation particulière du contour. En ce sens, notre critère n'impose aucune contrainte paramétrique au contour actif. Nous proposons deux dérivations possibles du critère d'a priori : par le calcul des variations, et par les gradients de forme.

Chapitre 6 : Applications. L'objectif de ce chapitre est de proposer différentes applications du critère d'a priori géométrique. Utilisé seul, ce critère est appliqué à la déformation de contours ou *shape warping*. Associé à des descripteurs statistiques de régions et des termes de régularisation, ce critère est appliqué à la segmentation et au suivi de régions homogènes.

La troisième partie est consacrée à la définition et l'étude d'un critère de **segmentation et estimation conjointes du mouvement**. Plusieurs applications de ce critère sont présentées. La troisième partie est composée des chapitres suivants :

Chapitre 7 : Introduction. Ce chapitre introduit notre approche à travers l'étude de l'état de l'art en détection et en segmentation du mouvement. Nous situons les principes et hypothèses de notre approche au vu des articles cités.

Chapitre 8 : Segmentation et estimation conjointes du mouvement à partir de deux images successives. Dans ce chapitre, nous définissons un critère de segmentation en mouvement conjointement à l'estimation du mouvement de la région segmentée. Nous dérivons le critère et proposons

deux équations d'évolution : une dirigée suivant la normale au contour, pour la segmentation d'objets en mouvement, et une dirigée suivant le mouvement de l'objet pour le suivi d'objets en mouvement. Nous proposons une méthode d'estimation du mouvement d'une région caractérisé par un modèle de mouvement.

Chapitre 9 : Segmentation et estimation conjointes du mouvement sur un groupe d'images. Ce chapitre étend la méthode de segmentation et d'estimation conjointes du mouvement du chapitre précédent. Le support temporel de la méthode s'étend de deux images à un groupe d'images. Nous développons les équations d'évolution du contour actif pour des applications de segmentation et de suivi d'objets en mouvement.

Chapitre 10 : Résultats expérimentaux. Ce chapitre présente les différents résultats obtenus avec le critère de segmentation et d'estimation conjointes du mouvement. Nous proposons d'évaluer la qualité de l'estimation du mouvement pour une région donnée en la comparant à des méthodes classiques d'estimation du flot optique, ou des méthodes de mise en correspondance. Puis, nous appliquons le critère conjoint à la segmentation d'images réelles. Enfin, nous présentons les résultats obtenus en suivi d'objets en mouvement en comparant les deux directions d'évolution du contour actif : suivant la normale au contour et suivant le mouvement.

Première partie

Introduction aux contours actifs

Chapitre 2

État de l'art des contours actifs

Un contour actif est une courbe qui évolue d'une forme initiale vers les frontières d'un objet d'intérêt, sous l'action d'une force (voir figure 2.1).

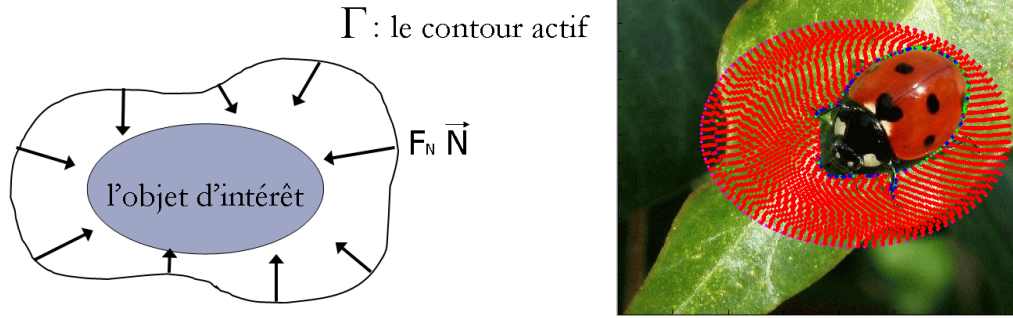


FIG. 2.1 – Principe des contours actifs : le contour actif Γ évolue vers l'objet d'intérêt sous l'action d'une force dirigée suivant la normale \mathbf{N} au contour.

Le contour actif Γ peut être défini comme une courbe paramétrée par $p \in [a, b]$ et de paramètre d'évolution $\tau \in [0, T]$ telle que :

$$\begin{aligned} \Gamma : [a, b] \times [0, T] &\rightarrow \mathbb{R}^2 \\ (p, \tau) &\mapsto \Gamma(p, \tau) = \mathbf{x}(p, \tau) = \begin{pmatrix} x(p, \tau) \\ y(p, \tau) \end{pmatrix} \end{aligned}$$

L'évolution du contour est régie par une équation de forme générale :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= \mathbf{F}(p, \tau) \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \tag{2.1}$$

où $\Gamma_0(p)$ est le contour initial.

Le contour se déforme suivant une force \mathbf{F} dont la direction est *a priori* quelconque.

Décomposons cette force dans un repère de Frenet lié au contour, dont la base est constituée du vecteur unitaire normal intérieur \mathbf{N} , et du vecteur unitaire tangent \mathbf{T} :

$$\mathbf{F} = F_N \mathbf{N} + F_T \mathbf{T} \quad (2.2)$$

La composante normale de la force, $F_N \mathbf{N}$, modifie la géométrie du contour, tandis que la composante tangentielle, $F_T \mathbf{T}$ n'influe que sur la paramétrisation du contour [28].

Par souci de simplicité, seule la composante normale est prise en compte, et l'équation d'évolution d'un contour actif s'écrit généralement :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= F_N(p, \tau) \mathbf{N}(p, \tau) \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (2.3)$$

Nous allons présenter quelques méthodes classiques de contours actifs : les snakes (en section 2.1), les contours actifs géométriques (en section 2.2), et les contours actifs géodésiques (en section 2.3). Si ces méthodes mettent en oeuvre des caractéristiques globales du contour, comme sa longueur ou sa rigidité, elles n'utilisent que des propriétés locales de l'image, comme son intensité ou son gradient en un point. Plus récemment, les méthodes de contours actifs s'attachent à caractériser les régions délimitées par le contour. Nous présenterons brièvement en section 2.4 quelques-unes de ces méthodes.

2.1 Snakes

Le premier modèle de contours actifs, appelé *snakes*, a été introduit par Kass *et al.* [52] en 1988. L'approche est variationnelle, c'est-à-dire que l'équation d'évolution du contour actif se déduit de la minimisation d'une énergie modélisant l'objet d'intérêt.

Le critère, ou l'énergie - nous emploierons indifféremment un des deux termes - est défini par Kass *et al.* comme :

$$J(\Gamma) = \alpha \int_a^b |\Gamma'(p)|^2 dp + \beta \int_a^b |\Gamma''(p)|^2 dp - \lambda \int_a^b |\nabla I(\Gamma(p))|^2 dp \quad (2.4)$$

où p paramétrise le contour $\Gamma : [a, b] \rightarrow \mathbb{R}^2$, et α, β et λ sont des constantes positives.

Les deux premiers termes définissent une contrainte interne au contour. Ce sont des termes de régularisation du contour qui déterminent son élasticité et sa rigidité.

Le dernier terme est un terme d'attache aux données. Il attire le contour vers les zones de forts gradients de l'image.

Cette approche est toutefois limitée par plusieurs inconvénients.

Le critère n'est pas intrinsèque, c'est-à-dire qu'il dépend de la paramétrisation du contour.

De plus, le contour initial doit être proche de l'objet pour que l'algorithme de minimisation converge.

La contrainte de régularité est forte et ne permet pas de détecter les zones concaves du contour. Elle interdit également les changements de topologie : un seul objet peut être segmenté.

Une telle régularité pousse le contour à rétrécir, voire à disparaître s'il ne rencontre pas de zones de gradient suffisamment fort. Ce sera le cas, par exemple, si le contour initial est à l'intérieur de l'objet à segmenter, et trop éloigné des frontières de l'objet.

Enfin, la composante de rigidité nécessite le calcul d'une dérivée à l'ordre quatre, ce qui entraîne des problèmes de discrétisation et d'instabilité numériques.

2.2 Contours actifs géométriques

Les approches géométriques, introduites par Osher *et al.* [67], puis Caselles *et al.* [9] et Malladi *et al.* [58] définissent une équation d'évolution qui ne provient plus de la minimisation d'un critère.

Ces approches utilisent une formulation par ensembles de niveaux [67].

Cette formulation permet la segmentation de plusieurs objets de même propriété à partir d'un seul contour initial.

Caselles *et al.* [9] proposent d'étendre l'équation de mouvement de la courbure moyenne :

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) |\nabla u| \quad (2.5)$$

Les auteurs ajoutent une constante réelle c , et introduisent la fonction de détection g :

$$\frac{\partial u}{\partial t} = g(I) \left(c + \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) \right) |\nabla u| \quad (2.6)$$

La fonction g est la fonction $g(I) = \frac{1}{1+(\nabla G_\sigma * I)^2}$, où le gradient de l'image I est convolué avec un filtre gaussien G_σ de variance σ .

C'est une attache aux données attirant le contour vers les frontières de l'objet et s'annulant aux frontières.

La constante c induit une force semblable à la force ballon introduite par Cohen [17]. Le paramètre c détermine le sens d'évolution du contour en fonction de son signe. Le contour, en l'absence de forts gradients, va rétrécir lorsque c est positif, et va pouvoir s'accroître, comme un ballon que l'on gonfle, lorsque c est négatif.

Les contraintes d'initialisation du contour actif sont, de ce fait, moins fortes. De plus, le contour peut segmenter des zones convexes.

L'équation d'évolution du contour actif s'écrit sous la forme :

$$\frac{\partial \Gamma}{\partial \tau} = g(I)(c + \kappa) \mathbf{N} \quad (2.7)$$

où est κ la courbure de Γ ($\kappa = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right)$ avec la formulation par ensembles de niveaux Γ).

2.3 Contours actifs géodésiques

Les contours actifs géodésiques reprennent l'approche variationnelle des snakes [52].

La fonctionnelle d'énergie est donnée par :

$$J(\Gamma) = \alpha \int_a^b |\Gamma'(p)|^2 dp + \lambda \int_a^b g^2(|\nabla I(\Gamma(p))|) dp \quad (2.8)$$

Par rapport à l'énergie définie par Kass *et al.* [52] (équation (2.4) page 12), on remarque l'abandon du terme de rigidité ($\beta = 0$), jugé par trop contraignant et provoquant des instabilités numériques.

On remarque de même l'introduction d'une fonction de détection g . La fonction $g : [0, +\infty] \rightarrow \mathbb{R}^+$ est une fonction strictement décroissante, telle que $\lim_{r \rightarrow +\infty} g(r) = 0$.

La fonction $g(r) = \frac{1}{1+r^m}$, $m = 1$ ou 2 est fréquemment choisie.

Dans [10], Caselles *et al.* proposent de minimiser l'énergie suivante :

$$J(\Gamma) = \int_0^{L(\Gamma)} g(|\nabla I(\Gamma(p))|) ds \quad (2.9)$$

avec $ds = |\Gamma'(p)| dp$.

Sur la notion d'équivalence entre les deux problèmes de minimisation voir [10] et [3].

Le problème s'interprète alors géométriquement comme la minimisation de la longueur du contour dans une métrique prenant en compte les caractéristiques de l'image.

La nouvelle fonctionnelle est intrinsèque, c'est-à-dire qu'elle ne dépend pas de la paramétrisation.

L'équation d'évolution déduite du critère (2.9) est :

$$\frac{\partial \Gamma}{\partial \tau} = (g(|\nabla I(\Gamma)|) \kappa - \nabla g(|\nabla I(\Gamma)|) \cdot \mathbf{N}) \mathbf{N} \quad (2.10)$$

où κ est la courbure du contour actif et \mathbf{N} le vecteur unitaire normal intérieur au contour.

Sur les zones homogènes, où le gradient de l'image est faible, le premier terme de l'équation d'évolution est prépondérant et le contour évolue suivant :

$$\frac{\partial \Gamma}{\partial \tau} = \kappa \mathbf{N} \quad (2.11)$$

Comme κ est une constante positive, et le vecteur \mathbf{N} est dirigé vers l'intérieur du contour, le contour actif aura donc tendance à rétrécir.

Sur les contours, où le gradient de l'image est de forte amplitude, le deuxième terme de l'équation d'évolution est prépondérant et le contour évolue suivant :

$$\frac{\partial \Gamma}{\partial \tau} = -\nabla g(|\nabla I(\Gamma)|) \cdot \mathbf{N} \mathbf{N}$$

Le contour actif s'approche de la frontière de l'objet, qu'il soit à l'intérieur ou à l'extérieur de l'objet (voir la figure 2.2).

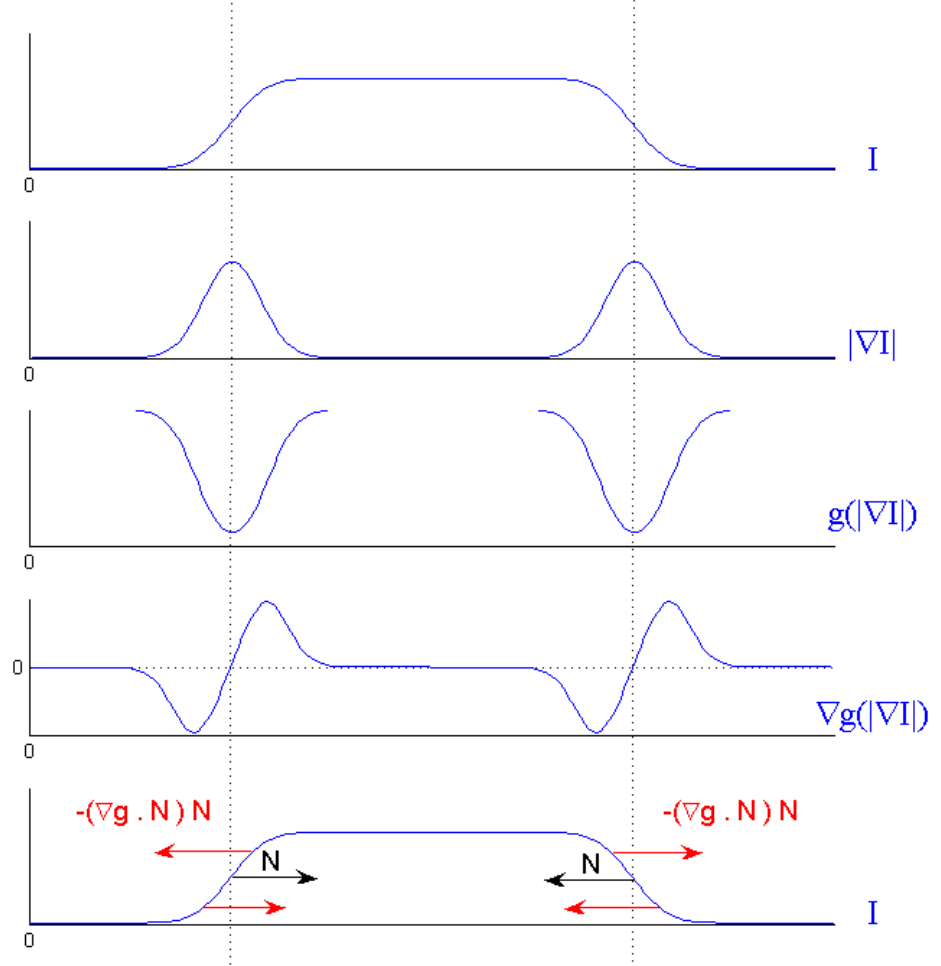


FIG. 2.2 – Évolution du contour actif suivant l'équation $\frac{\partial \Gamma}{\partial \tau} = -\nabla g(|\nabla I(\Gamma)|) \cdot \mathbf{N}\mathbf{N}$ près des frontières de l'objet.

2.4 Descripteurs de région

Toutes ces approches s'attachent à décrire le contour à partir de données locales liées à l'image, en particulier son gradient, et se traduisent par des intégrales sur le contour. Les seuls termes globaux sont ceux liés au contour, comme ceux qui minimisent sa longueur.

Afin de pouvoir prendre en compte des propriétés intrinsèques à l'objet, comme sa moyenne, sa texture ou son mouvement, l'énergie doit contenir des intégrales de région :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x}. \quad (2.12)$$

La fonction $k(\mathbf{x}, \Omega)$ est appelée descripteur de région.

Les premiers contours actifs utilisant des descripteurs de région ont été introduits par Cohen *et al.* [17] et Ronfard [74].

Cohen *et al.* [17] utilisent une approche variationnelle de contours actifs pour la reconstruction de surfaces. Ils définissent une énergie caractérisant deux surfaces à reconstruire, avec des propriétés différentes, et le contour, frontière de ces surfaces.

Ronfard [74] propose de segmenter l'image en deux régions homogènes à l'aide de descripteurs statistiques k^{in} et k^{out} . Le contour évolue dans la direction de la normale suivant une force proportionnelle à la différence des descripteurs : $k^{in} - k^{out}$.

Dans le cadre d'une segmentation de l'image en deux régions, Zhu et Yuille [93] proposent un algorithme de "compétition de régions" combinant les propriétés géométriques des contours actifs aux techniques statistiques de croissance de région.

Chan et Vese [12] s'inspirent des travaux de Mumford et Shah [63] pour proposer un algorithme de segmentation en deux régions, s'appuyant sur leur moyenne, qu'ils formulent selon la méthode des ensembles de niveaux.

Paragios *et al.* [70] étendent l'approche des contours géodésiques en incluant des termes de région.

Galland *et al.* [31] proposent un algorithme de segmentation en zones homogènes d'images fortement bruitées, à l'aide d'une grille active polygonale, et basée sur la minimisation de la complexité stochastique de l'image.

Debreuve *et al.* [23, 24] proposent un critère spatio-temporel pour la segmentation de séquences cardiaques. Ils développent parallèlement une méthode géométrique de recalage, non rigide et non paramétrique, utilisant

les propriétés de la méthode des ensembles de niveaux.

Jehan *et al.* [48, 51] différentient le critère à l'aide d'outils inspirés de l'optimisation de domaines. Les auteurs montrent que, lorsque le descripteur de la région dépend lui-même de la région, il induit des termes supplémentaires dans l'équation d'évolution.

Precioso *et al.* [71, 72] proposent une mise en œuvre paramétrique des contours actifs, en étendant le modèle classique des Splines d'interpolation à un modèle de Splines d'approximation, les *smoothing splines*. Ces splines ont des propriétés intrinsèques de régularisation favorisant la robustesse au bruit et leur utilisation permet de réduire considérablement le temps de calcul par rapport aux ensembles de niveaux.

Herbulot *et al.* [37] définissent des critères statistiques appliqués à des données multimodales en imagerie médicale. L'approche variationnelle mise en œuvre utilise les gradients de forme.

Approche variationnelle : dérivation par les gradients de forme

Notre approche est une approche variationnelle c'est-à-dire que le problème de segmentation est formulé comme la minimisation d'un critère. Il s'agit de trouver le domaine qui minimise le critère définissant l'objet d'intérêt.

L'introduction d'un schéma dynamique permet de dériver le critère par rapport à un paramètre d'évolution du domaine. La dérivation du critère peut alors s'effectuer en utilisant des outils de dérivation classique, comme le calcul des variations, ou en utilisant des outils issus de l'optimisation de domaine : les gradients de forme. L'équation d'évolution du contour actif est déduite de la dérivée du critère. La figure 3.1 page 20 résume les principales étapes de l'approche variationnelle avec calcul de l'équation d'évolution pour un terme de région.

Dans ce chapitre, nous définirons les types de termes pouvant constituer un critère, à savoir : termes de contour et termes de région (section 3.1 page 21).

Jehan *et al.* [48, 49, 51] mettent l'accent sur le fait que certains descripteurs de région, comme la moyenne, la variance ou les histogrammes apportent des informations globales sur les propriétés des régions. Les termes de région du critère dépendent alors doublement de la région : comme intégrale sur la région d'un descripteur dépendant lui-même de la région. On dira ces termes à *double dépendance*.

Nous présenterons ensuite les outils de dérivation de domaine mis en œuvre par Jehan *et al.* dans [48] (section 3.2 page 22).

L'originalité principale de leur travail réside dans l'utilisation d'outils de

dérivation issus de l'optimisation de domaine, et dans la prise en compte de cette double dépendance de certains termes de région. Les auteurs montrent également l'importance des termes additifs qui en découlent.

Enfin, nous dériverons chacun des termes de contour et de région (section 3.3 page 25). Le cas des termes de région à double dépendance sera introduit à l'aide d'exemples concrets afin de mieux appréhender le cadre général présenté en section 3.4.2 page 35 .

Définition du critère

- ▷ En fonction de l'application
- ▷ Termes de contour ou de région

$$\textit{Exemple d'un terme de région} : J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) \, d\mathbf{x}$$

Introduction d'un schéma dynamique

- ▷ Famille de transformations

$$J(\Omega(\tau)) = \int_{\Omega(\tau)} k(\mathbf{x}, \Omega(\tau)) \, d\mathbf{x}$$

Dérivation du critère

- ▷ Outils de dérivation classiques
- ▷ Méthode des gradients de forme

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} F(s, \Omega) \mathbf{V} \cdot \mathbf{N} \, ds$$

Équation d'évolution

- ▷ Direction : suivant la normale
- ▷ Assurer une dérivée négative

$$\frac{\partial \Gamma}{\partial \tau}(p, \tau) = F(p, \tau) \, \mathbf{N}$$

$$\Gamma(p, 0) = \Gamma_0(p)$$

FIG. 3.1 – Approche variationnelle avec calcul de l'EDP.

3.1 Descripteurs de contour et de région

Une des premières difficultés d'une approche variationnelle est de formuler le problème à résoudre comme la minimisation d'un critère.

Ce critère peut être composé de termes traduisant des informations sur le contour de la région, comme par exemple le gradient de l'image aux points du contour, et/ou de termes caractérisant les propriétés intrinsèques de la région, comme sa moyenne, sa variance, son histogramme ou sa vitesse.

De manière générale, un terme de contour s'écrit :

$$J(\Gamma) = \int_{\Gamma} k(s) ds \quad (3.1)$$

où $k(s)$ est un descripteur du contour Γ , contour paramétré par son abscisse curviligne s .

Par exemple, nous développerons, dans le chapitre II, le cas d'un descripteur contour, fonction de la distance du contour actif à un contour de référence.

Un terme de région, quant à lui, s'écrit de manière générale :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x} \quad (3.2)$$

où $k(\mathbf{x}, \Omega)$ est un descripteur de région.

Les descripteurs de région peuvent dépendre ou non de la région qu'ils caractérisent.

Par exemple, le descripteur de région $k = \alpha$, où α est une constante positive, est un descripteur indépendant de la région d'intégration. Le terme de région associé est un terme de régularisation qui tend à réduire l'aire de la région (l'objet en l'occurrence).

Autre exemple, le descripteur de région $k = |I_n - I_{n-1}^{comp}|$ calcule la différence entre l'intensité du fond de l'image I_n et du fond, compensé en mouvement, de l'image précédente I_{n-1} . Un tel critère sert pour des applications de détection d'objets en mouvement. Utilisé seul il tend à sur-segmenter l'objet, c'est-à-dire que des pixels du fond de l'image sont considérés comme appartenant à l'objet (Annexe V page 131).

Cependant la plupart des descripteurs, notamment les descripteurs statistiques, dépendent de la région qu'ils caractérisent.

Par exemple, Chan *et al.* [12] et Debreuve *et al.* [24] utilisent comme descripteur de région homogène $k = (I - \mu)^2$, où μ est la moyenne de l'intensité de l'image I sur la région étudiée.

Jehan *et al.* [48] utilisent comme descripteurs, pour segmenter des images couleurs en régions homogènes, des fonctions de la variance comme le déterminant de la matrice de covariance, mais aussi des histogrammes.

Tous ces descripteurs dépendent de la région et induisent une double dépendance du terme de région par rapport à la région.

3.2 Outils de dérivation

Dans cette section, nous présenterons les outils de dérivation mis en oeuvre pour le calcul de l'équation d'évolution correspondant à chaque type de terme de contour ou de région.

Tout d'abord, nous introduirons un schéma dynamique, afin de pouvoir dériver le critère par rapport au paramètre d'évolution du contour.

Puis, nous énoncerons des définitions de dérivées, dont la dérivée de domaine, et le théorème d'équivalence entre dérivée eulérienne et dérivée de domaine.

Ensuite, nous étudierons la dérivation de critères introduits dans un ordre de complexité croissant : nous commencerons par un terme de contour, puis un terme de région tous deux indépendants du domaine d'intégration ; ensuite, nous introduirons, par quelques exemples concrets, la dérivation des termes de région à double dépendance, afin d'illustrer la méthode des gradients de forme.

Enfin, nous étendrons cette méthode à un cadre général par l'énoncé d'un théorème donnant la dérivée d'un terme de région à double dépendance.

3.2.1 Introduction du schéma dynamique dans le critère

Un critère région s'écrit généralement comme l'intégrale, sur la région considérée, d'un descripteur de cette région, soit :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x} \quad (3.3)$$

De même, un critère contour s'écrit généralement comme l'intégrale sur le contour d'un descripteur de contour, soit :

$$J(\Gamma) = \int_{\Gamma} k(s) ds \quad (3.4)$$

L'ensemble des domaines de \mathcal{R}^n n'a pas une structure d'espace vectoriel. On ne peut donc pas dériver le critère par rapport au domaine Ω (respectivement Γ).

Les techniques de dérivation de domaines [26, 38] introduisent une famille de transformations T_τ , appartenant à un espace vectoriel, qui permettent de faire évoluer le domaine Ω (respectivement Γ) en fonction de τ de sorte que :

$$\begin{aligned} T_\tau : \Omega &\rightarrow \Omega(\tau) \quad \text{et} \quad T_0(\Omega) = \Omega \\ (\text{respectivement } T_\tau : \Gamma &\rightarrow \Gamma(\tau) \quad \text{et} \quad T_0(\Gamma) = \Gamma) \end{aligned} \quad (3.5)$$

Dans [26], Sokolowsky et Zolésio définissent la famille de transformations à partir de la vitesse de déformation des domaines. Nous reprenons ici leur raisonnement en l'adaptant à notre notation.

Si l'on considère que le domaine Ω se déforme suivant un paramètre $\tau \in [0, \epsilon[$, qui représente par exemple le temps, tout point \mathbf{x} de Ω est transformé, au temps τ suivant $x(\tau, \mathbf{x})$.

La vitesse eulérienne de ce point est : $\mathbf{V}(\tau, x(\tau, \mathbf{x})) = \frac{\partial}{\partial \tau} x(\tau, \mathbf{x})$. Cette vitesse est considérée comme la direction de déformation et permet de reconstruire la déformation.

Soit $V(\tau, x(\tau, \mathbf{x})) \in \mathcal{C}^0([0, \epsilon], \mathcal{C}^1(\mathcal{R}^n, \mathcal{R}^n))$. On associe à ce champ l'équation différentielle :

$$\begin{aligned} \frac{\partial}{\partial \tau} x(\tau, \mathbf{x}) &= V(\tau, x(\tau, \mathbf{x})) \\ x(0, \mathbf{x}) &= \mathbf{x} \end{aligned} \quad (3.6)$$

Si le champ de vitesse V est borné sur \mathcal{R}^n , la solution $x(\tau, \mathbf{x})$ est définie pour $\tau \in [0, \epsilon[$ et $\mathbf{x} \in \mathcal{R}^2$, et on définit l'application $T_\tau : \mathbf{x} \rightarrow x(\tau, \mathbf{x})$.

Le domaine déformé $\Omega(\tau)$ est $\Omega(\tau) = T_\tau(\Omega)$ soit :
 $\Omega(\tau) = \{x \in \mathcal{R}^n / x = x(\tau, \mathbf{x}), \forall \mathbf{x} \in \Omega\}$

Nous introduisons ainsi un schéma dynamique dans le critère :

$$J(\Omega(\tau)) = \int_{\Omega(\tau)} k(x, \Omega(\tau)) \, dx \quad (3.7)$$

et pour un critère contour :

$$J(\Gamma(\tau)) = \int_{\Gamma(\tau)} k(s) \, ds \quad (3.8)$$

3.2.2 Définition d'une dérivée de domaine et théorème d'équivalence

Nous allons tout d'abord rappeler les définitions de la dérivée eulérienne et de la dérivée de domaine. Puis nous présenterons un théorème, maintes fois utilisé par la suite, donnant l'expression de la dérivée eulérienne en fonction de la dérivée de domaine.

Lors des calculs des dérivées, nous supposons des petites déformations des domaines, le développement limité au premier ordre de la transformation s'écrit :

$$T_\tau(\Omega) = T_0(\Omega) + \tau \frac{\partial T_0}{\partial \tau}(\Omega) \quad (3.9)$$

Comme $T_0(\Omega) = \Omega$ (équation (3.5)) et

$$\begin{aligned} \mathbf{V}(\tau, \Omega(\tau)) &= \frac{\partial}{\partial \tau}(\Omega(\tau)) = \frac{\partial T_\tau}{\partial \tau}(\Omega) \\ \mathbf{V}(0, \Omega(0)) &= \frac{\partial T_0}{\partial \tau}(\Omega) \stackrel{\text{notation}}{=} \mathbf{V} \end{aligned} \quad (3.10)$$

On écrit alors :

$$T_\tau(\Omega) = \Omega(\tau) = \Omega + \tau \mathbf{V} \quad (3.11)$$

Rappelons pour commencer la définition de la dérivée eulérienne :

Définition 3.1. *La dérivée eulérienne de la fonctionnelle de domaine $J(\Omega) = \int_\Omega k(\mathbf{x}, \Omega) d\mathbf{x}$ dans la direction du champ de vecteur \mathbf{V} est définie comme la limite :*

$$dJ(\Omega, \mathbf{V}) = \lim_{\tau \rightarrow 0} \frac{J(\Omega + \tau \mathbf{V}) - J(\Omega)}{\tau}$$

Définition 3.2. *La dérivée matérielle de $k(\mathbf{x}, \Omega)$ dans la direction \mathbf{V} , notée $\dot{k}(\mathbf{x}, \Omega, \mathbf{V})$ est définie par :*

$$\dot{k}(\mathbf{x}, \Omega, \mathbf{V}) = \lim_{\tau \rightarrow 0} \frac{k(\mathbf{x} + \tau \mathbf{V}, \Omega + \tau \mathbf{V}) - k(\mathbf{x}, \Omega)}{\tau}$$

La dérivée de domaine est définie comme suit :

Définition 3.3. *La dérivée de domaine de $k(\mathbf{x}, \Omega)$ dans la direction \mathbf{V} , notée $d_\Omega k(\mathbf{x}, \Omega, \mathbf{V})$ est définie par :*

$$d_\Omega k(\mathbf{x}, \Omega, \mathbf{V}) = \lim_{\tau \rightarrow 0} \frac{k(\mathbf{x}, \Omega + \tau \mathbf{V}) - k(\mathbf{x}, \Omega)}{\tau}$$

Le théorème suivant établit une relation entre la dérivée eulérienne et la dérivée de domaine pour une intégrale de région $J(\Omega)$:

Théorème 3.1. *La dérivée eulérienne $dJ(\Omega, \mathbf{V})$ de la fonctionnelle $J(\Omega) = \int_\Omega k(\mathbf{x}, \Omega) d\mathbf{x}$ dans la direction \mathbf{V} est la suivante :*

$$dJ(\Omega, \mathbf{V}) = \int_\Omega d_\Omega k(\mathbf{x}, \Omega, \mathbf{V}) d\mathbf{x} - \int_\Gamma k(s, \Omega) \mathbf{V} \cdot \mathbf{N} ds$$

où \mathbf{N} est la normale unitaire intérieure à Γ et s son abscisse curviligne.

Une démonstration de ce théorème peut être trouvée dans [26, 80].

Ce théorème, utilisé plusieurs fois si nécessaire, nous permettra, dans des cas favorables, de convertir les intégrales de région en une expression équivalente basée sur une intégrale de contour (voir section 3.4.2 page 35). L'équation d'évolution du contour actif se déduit aisément d'une telle expression de la dérivée du critère.

3.3 Dérivation de termes de contour et de région

Nous allons tout d'abord étudier la dérivation d'un terme de contour indépendant du domaine d'intégration.

Ensuite nous verrons le cas d'un terme de région dont le descripteur est indépendant de la région.

3.3.1 Dérivation d'un terme de contour

Un terme de contour a pour forme générale :

$$J(\Gamma) = \int_\Gamma k(s) ds \tag{3.12}$$

Sa dérivée est bien connue et calculée entre autres par Caselles *et al* [10] :

Théorème 3.2. *La dérivée eulérienne $dJ(\Gamma, \mathbf{V})$ d'un terme de contour de la forme $J(\Gamma) = \int_{\Gamma} k(s) ds$ dans la direction \mathbf{V} est :*

$$dJ(\Gamma, \mathbf{V}) = \int_{\Gamma} (\nabla k(s) \cdot \mathbf{N} - k(s) \kappa) \mathbf{V} \cdot \mathbf{N} ds$$

où \mathbf{N} est la normale unitaire intérieure de Γ , κ la courbure moyenne et s l'abscisse curviligne de Γ .

L'équation d'évolution qui en découle est donnée par :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= (k(\mathbf{s}) \kappa - \nabla k(\mathbf{s}) \cdot \mathbf{N}) \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.13)$$

où $\Gamma_0(p)$ est le contour initial, $\mathbf{s} = \Gamma(p, \tau)$ et p paramétrise le contour.

L'équation d'évolution se déduit de la dérivée eulérienne : on assure la minimisation du critère en faisant évoluer le contour actif dans la direction \mathbf{N} , et en choisissant une vitesse \mathbf{V} telle que la dérivée du critère soit négative.

La vitesse du contour actif est :

$$\mathbf{V} = (k(\mathbf{s}) \kappa - \nabla k(\mathbf{s}) \cdot \mathbf{N}) \mathbf{N}, \text{ avec } \mathbf{s} = \Gamma(p, \tau) \quad (3.14)$$

et l'équation d'évolution qui en résulte est :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= (k(\mathbf{s}) \kappa - \nabla k(\mathbf{s}) \cdot \mathbf{N}) \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.15)$$

où $\Gamma_0(p)$ est un contour initial défini par l'utilisateur, $\mathbf{s} = \Gamma(p, \tau)$ et p paramétrise le contour.

3.3.2 Dérivation d'un terme de région dont le descripteur ne dépend pas de la région

Un terme de région dont le descripteur ne dépend pas de la région s'écrit de manière générale :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}) d\mathbf{x} \quad (3.16)$$

Corollaire 3.1. *La dérivée eulérienne $dJ(\Omega, \mathbf{V})$ d'un terme de région de la forme $J(\Omega) = \int_{\Omega} k(\mathbf{x}) d\mathbf{x}$ dans la direction \mathbf{V} est :*

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} k(s) \mathbf{V} \cdot \mathbf{N} ds \quad (3.17)$$

où \mathbf{N} est la normale unitaire intérieure de Γ et s l'abscisse curviligne.

L'équation d'évolution qui en découle est donnée par :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= k(\mathbf{s}) \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.18)$$

où $\Gamma_0(p)$ est un contour initial défini par l'utilisateur, $\mathbf{s} = \Gamma(p, \tau)$ et p paramétrise le contour.

Démonstration :

La dérivation d'un terme de région dont le descripteur ne dépend pas de la région est simple car elle utilise directement le théorème 3.1.

En effet, comme le descripteur ne dépend pas de la région Ω :

$$k(\mathbf{x}, \Omega(\tau)) = k(\mathbf{x}, \Omega) = k(\mathbf{x})$$

et la dérivée de domaine est nulle :

$$d_{\Omega} k(\mathbf{x}, \Omega, \mathbf{V}) = 0 \quad (3.19)$$

Ainsi la dérivée eulérienne de J se réduit à l'expression suivante :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} k(s) \mathbf{V} \cdot \mathbf{N} ds \quad (3.20)$$

L'équation d'évolution se déduit de la dérivée eulérienne en faisant évoluer le contour actif dans la direction \mathbf{N} , et en choisissant une vitesse \mathbf{V} telle que la dérivée du critère soit négative.

L'équation d'évolution du contour actif qui en découle est :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= k(\mathbf{s}) \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.21)$$

où $\Gamma_0(p)$ est un contour initial défini par l'utilisateur, $\mathbf{s} = \Gamma(p, \tau)$ et p paramétrise le contour.

□

3.4 Dérivation d'un terme de région à double dépendance

Avant de présenter le théorème général qui permet de calculer la dérivée d'un terme de région lorsque son descripteur dépend de la région, nous allons introduire, par quelques exemples concrets, la méthodologie qui sert de fondement à l'élaboration de ce théorème.

3.4.1 Exemples de dérivation d'un terme de région à double dépendance

Nous allons étudier deux descripteurs de régions dépendants de la région qu'ils décrivent. L'un est fonction de la moyenne de la région, l'autre est fonction de la variance de la région. Ces deux cas particuliers induisent une double dépendance du terme région par rapport à la région. Ces exemples devraient permettre de mieux appréhender la méthodologie sous-jacente à l'élaboration du théorème 3.3 page 36.

Descripteur fonction de la moyenne

Un terme de région dont le descripteur est fonction de la moyenne $\mu(\Omega)$ de l'intensité $I(\mathbf{x})$ d'une image sur la région Ω peut s'écrire :

$$J(\Omega) = \int_{\Omega} \varphi(I(\mathbf{x}) - \mu(\Omega)) \, d\mathbf{x} \quad (3.22)$$

où φ est une fonction dérivable, en général paire et croissante, mais seule la dérivabilité de la fonction est nécessaire au calcul de l'équation d'évolution. Un tel critère sert à segmenter les régions homogènes, car on cherche la région dont les pixels ont une intensité proche de la moyenne de la région.

On calcule la dérivée eulérienne du critère ainsi défini en utilisant le théorème 3.1. Ce théorème, appliqué récursivement, va nous permettre d'obtenir finalement une expression du critère basée sur une intégrale de contour.

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} d_{\Omega}(\varphi(I(\mathbf{x}) - \mu(\Omega)))(\mathbf{x}, \Omega, \mathbf{V}) \, d\mathbf{x} - \int_{\Gamma} \varphi(I(s) - \mu(\Omega)) \mathbf{V} \cdot \mathbf{N} \, ds \quad (3.23)$$

où la dérivée de domaine dans la direction \mathbf{V} de la fonction composée $\varphi(I(\mathbf{x}) - \mu(\Omega))$ a pour expression :

$$\begin{aligned} d_{\Omega}(\varphi(I(\mathbf{x}) - \mu(\Omega)))(\mathbf{x}, \Omega, \mathbf{V}) &= -d_{\Omega}\mu(\mathbf{x}, \Omega, \mathbf{V}) \, \varphi'(I(\mathbf{x}) - \mu(\Omega)) \\ &= -d\mu(\Omega, \mathbf{V}) \, \varphi'(I(\mathbf{x}) - \mu(\Omega)) \end{aligned} \quad (3.24)$$

où $\varphi'(r) = \frac{d\varphi(r)}{dr}$.

Calculons donc la dérivée eulérienne de $\mu(\Omega)$ dans la direction \mathbf{V} .

La fonction $\mu(\Omega)$ peut s'écrire comme fonction composée de fonctions de Ω :

$$\mu(\Omega) = \frac{1}{|\Omega|} \int_{\Omega} I(\mathbf{x}) \, d\mathbf{x} = f(G_1(\Omega), G_2(\Omega)) \quad (3.25)$$

avec $G_1 = \int_{\Omega} I(\mathbf{x}) \, d\mathbf{x}$ et $G_2 = \int_{\Omega} d\mathbf{x} \stackrel{\text{notation}}{=} |\Omega|$.

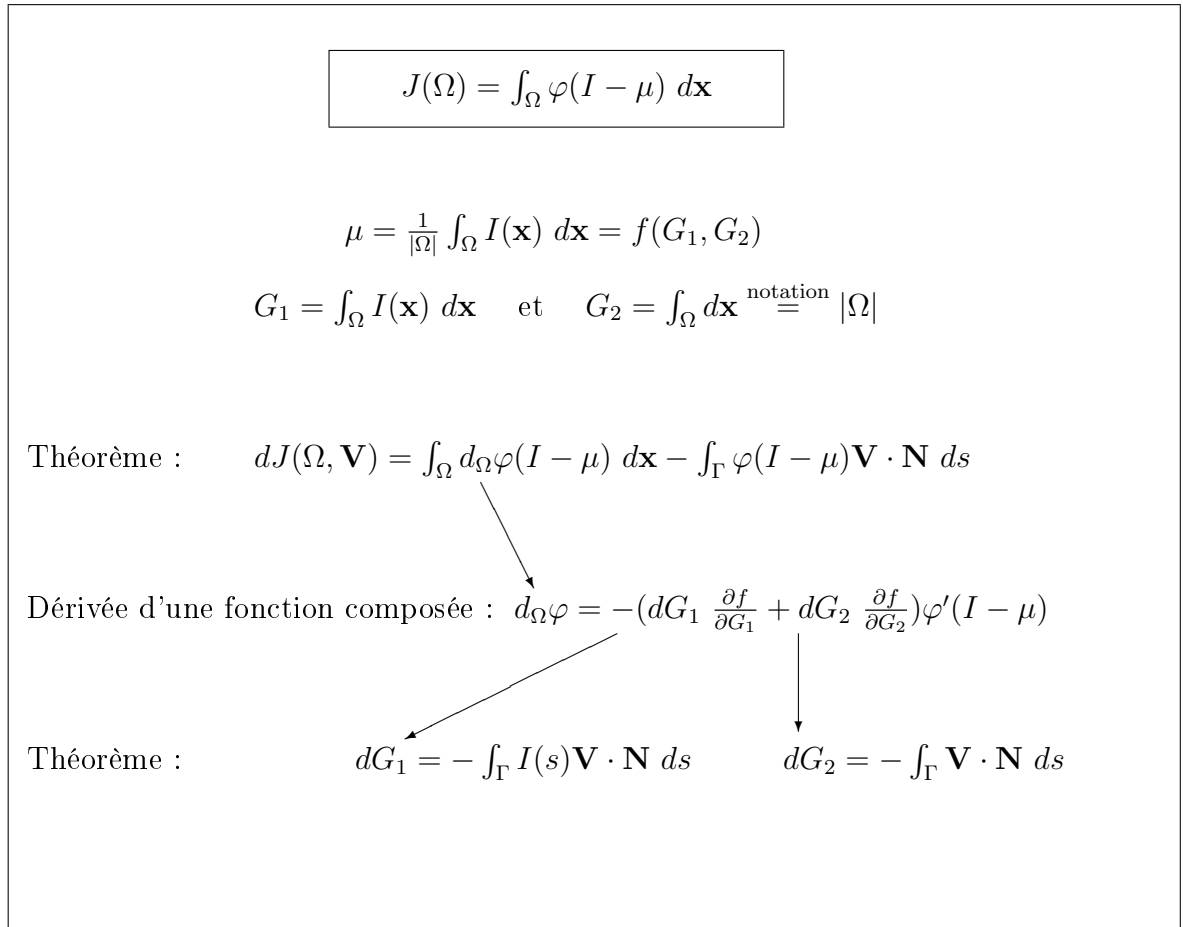


FIG. 3.2 – Étapes de dérivation d'un terme de région dont le descripteur est une fonction de la moyenne de la région.

La dérivée eulérienne de f dans la direction \mathbf{V} se décompose suivant :

$$df(\Omega, \mathbf{V}) = dG_1(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_1}(G_1, G_2) + dG_2(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_2}(G_1, G_2) \quad (3.26)$$

Le calcul des dérivées partielles de f par rapport à G_1 et G_2 est direct et donne :

$$\frac{\partial f}{\partial G_1} = \frac{1}{G_2} = \frac{1}{|\Omega|} \quad (3.27)$$

$$\frac{\partial f}{\partial G_2} = -\frac{G_1}{G_2^2} = -\frac{\mu(\Omega)}{|\Omega|} \quad (3.28)$$

Il reste donc à calculer les deux dérivées eulériennes de G_1 et G_2 dans la direction \mathbf{V} en utilisant à nouveau le théorème 3.1 :

$$G_1 = \int_{\Omega} I(\mathbf{x}) \, d\mathbf{x}$$

et sa dérivée eulérienne vaut :

$$dG_1(\Omega, \mathbf{V}) = \int_{\Omega} d_{\Omega} I(\mathbf{x}, \Omega, \mathbf{V}) \, d\mathbf{x} - \int_{\Gamma} I(s) \mathbf{V} \cdot \mathbf{N} \, ds = - \int_{\Gamma} I(s) \mathbf{V} \cdot \mathbf{N} \, ds \quad (3.29)$$

car, comme $I(\mathbf{x})$ est constante par rapport à τ , $d_{\Omega} I(\mathbf{x}, \Omega, \mathbf{V}) = 0$.

La dérivée de $G_2 = \int_{\Omega} d\mathbf{x}$ est donnée par :

$$dG_2(\Omega, \mathbf{V}) = \int_{\Omega} d_{\Omega} \mathbf{1}(\mathbf{x}, \Omega, \mathbf{V}) \, d\mathbf{x} - \int_{\Gamma} \mathbf{V} \cdot \mathbf{N} \, ds = - \int_{\Gamma} \mathbf{V} \cdot \mathbf{N} \, ds \quad (3.30)$$

où $\mathbf{1}$ représente la fonction constante égale à 1.

Des équations (3.27), (3.28), (3.29) et (3.30) l'on déduit l'expression de la dérivée eulérienne de μ dans la direction \mathbf{V} :

$$d\mu(\Omega, \mathbf{V}) = - \int_{\Gamma} \frac{I(s) - \mu(\Omega)}{|\Omega|} \mathbf{V} \cdot \mathbf{N} \, ds \quad (3.31)$$

Enfin, en remplaçant dans (3.23) $d_{\Omega} \varphi$ par son expression dans (3.24), et $d\mu(\Omega, \mathbf{V})$ dans (3.24) par son expression dans (3.31), on obtient :

$$\begin{aligned} dJ(\Omega, \mathbf{V}) &= \int_{\Omega} \varphi'(I(\mathbf{x}) - \mu(\Omega)) \int_{\Gamma} \frac{I(s) - \mu(\Omega)}{|\Omega|} \mathbf{V} \cdot \mathbf{N} \, ds \, d\mathbf{x} \\ &\quad - \int_{\Gamma} \varphi(I(s) - \mu(\Omega)) \mathbf{V} \cdot \mathbf{N} \, ds \end{aligned} \quad (3.32)$$

soit encore, en permutant les intégrales :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} \left[\varphi(I(s) - \mu(\Omega)) - \frac{I(s) - \mu(\Omega)}{|\Omega|} \int_{\Omega} \varphi'(I(\mathbf{x}) - \mu(\Omega)) \, d\mathbf{x} \right] \mathbf{V} \cdot \mathbf{N} \, ds \quad (3.33)$$

L'équation d'évolution associée est :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= \left[\varphi(I(\mathbf{s}) - \mu(\Omega(\tau))) - \frac{I(\mathbf{s}) - \mu(\Omega(\tau))}{|\Omega(\tau)|} \int_{\Omega(\tau)} \varphi'(I(\mathbf{x}) - \mu(\Omega(\tau))) d\mathbf{x} \right] \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.34)$$

avec $\mathbf{s} = \Gamma(p, \tau)$.

Intéressons nous plus précisément au cas particulier de la fonction $\varphi : \varphi(r) = r^2$. Nous utiliserons ces résultats au paragraphe suivant pour le calcul de la dérivée d'un descripteur fonction de la variance.

Si la fonction de pondération φ est la fonction carrée, le critère s'écrit :

$$J(\Omega) = \int_{\Omega} (I(\mathbf{x}) - \mu(\Omega))^2 d\mathbf{x} \quad (3.35)$$

En adaptant à la fonction $\varphi(r) = r^2$ l'équation de la dérivée de J pour le cas général (3.33) on obtient :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} \left[(I(s) - \mu(\Omega))^2 - \frac{I(s) - \mu(\Omega)}{|\Omega|} \int_{\Omega} 2(I(\mathbf{x}) - \mu(\Omega)) d\mathbf{x} \right] \mathbf{V} \cdot \mathbf{N} ds \quad (3.36)$$

or

$$\int_{\Omega} I(\mathbf{x}) - \mu(\Omega) = 0$$

ce qui simplifie l'expression de la dérivée du critère :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} (I(s) - \mu(\Omega))^2 \mathbf{V} \cdot \mathbf{N} ds \quad (3.37)$$

L'équation d'évolution associée est :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= [I(\mathbf{s}) - \mu(\Omega(\tau))]^2 \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.38)$$

avec $\mathbf{s} = \Gamma(p, \tau)$.

Descripteur fonction de la variance

L'exemple d'un descripteur de région fonction de la variance est un peu plus complexe que l'exemple précédent, car il nécessite une décomposition supplémentaire du descripteur en fonctions de la région. La figure 3.3 illustre les

différentes étapes nécessaires à la dérivation de ce terme de région.

$$J(\Omega) = \int_{\Omega} \varphi(\sigma^2(\Omega)) \, d\mathbf{x} \quad (3.39)$$

où $\sigma^2(\Omega)$ représente la variance de la région Ω .

Appliquons le théorème 3.1 pour exprimer la dérivée eulérienne du critère comme la somme de l'intégrale sur la région de sa dérivée de domaine et d'une intégrale de contour.

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} d_{\Omega}(\varphi(\sigma^2))(\mathbf{x}, \Omega, \mathbf{V}) d\mathbf{x} - \int_{\Gamma} \varphi(\sigma^2(\Omega)) \, \mathbf{V} \cdot \mathbf{N} \, ds \quad (3.40)$$

Considérons la dérivée de domaine $d_{\Omega}(\varphi(\sigma^2))(\mathbf{x}, \Omega, \mathbf{V})$ comme une dérivée de fonctions composées :

$$d_{\Omega}\varphi(\sigma^2)(\mathbf{x}, \Omega, \mathbf{V}) = d\sigma^2(\Omega, \mathbf{V})\varphi'(\sigma^2(\Omega)) \quad (3.41)$$

où $\varphi'(r) = \frac{d\varphi(r)}{dr}$.

La variance peut s'écrire comme une fonction composée de deux fonctions de Ω :

$$\sigma^2(\Omega) = \frac{1}{|\Omega|} \int_{\Omega} (I(\mathbf{x}) - \mu(\Omega))^2 \, d\mathbf{x} = f(G_1, G_2) \quad (3.42)$$

avec $G_1(\Omega) = \int_{\Omega} (I(\mathbf{x}) - \mu(\Omega))^2 \, d\mathbf{x}$ et $G_2(\Omega) = \int_{\Omega} d\mathbf{x}$.

Sa dérivée eulérienne dans la direction \mathbf{V} a pour expression :

$$d\sigma^2(\Omega, \mathbf{V}) = dG_1(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_1}(G_1, G_2) + dG_2(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_2}(G_1, G_2) \quad (3.43)$$

Les calculs de $\frac{\partial f}{\partial G_1}$ et $\frac{\partial f}{\partial G_2}$ sont directs :

$$\frac{\partial f}{\partial G_1}(G_1, G_2) = \frac{1}{|\Omega|} \quad (3.44)$$

et

$$\frac{\partial f}{\partial G_2}(G_1, G_2) = -\frac{G_1}{G_2^2} = -\frac{\sigma^2}{|\Omega|} \quad (3.45)$$

Pour calculer $dG_1(\Omega, \mathbf{V})$ nous utilisons le résultat démontré au paragraphe précédent, pour le cas $\varphi(r) = r^2$.

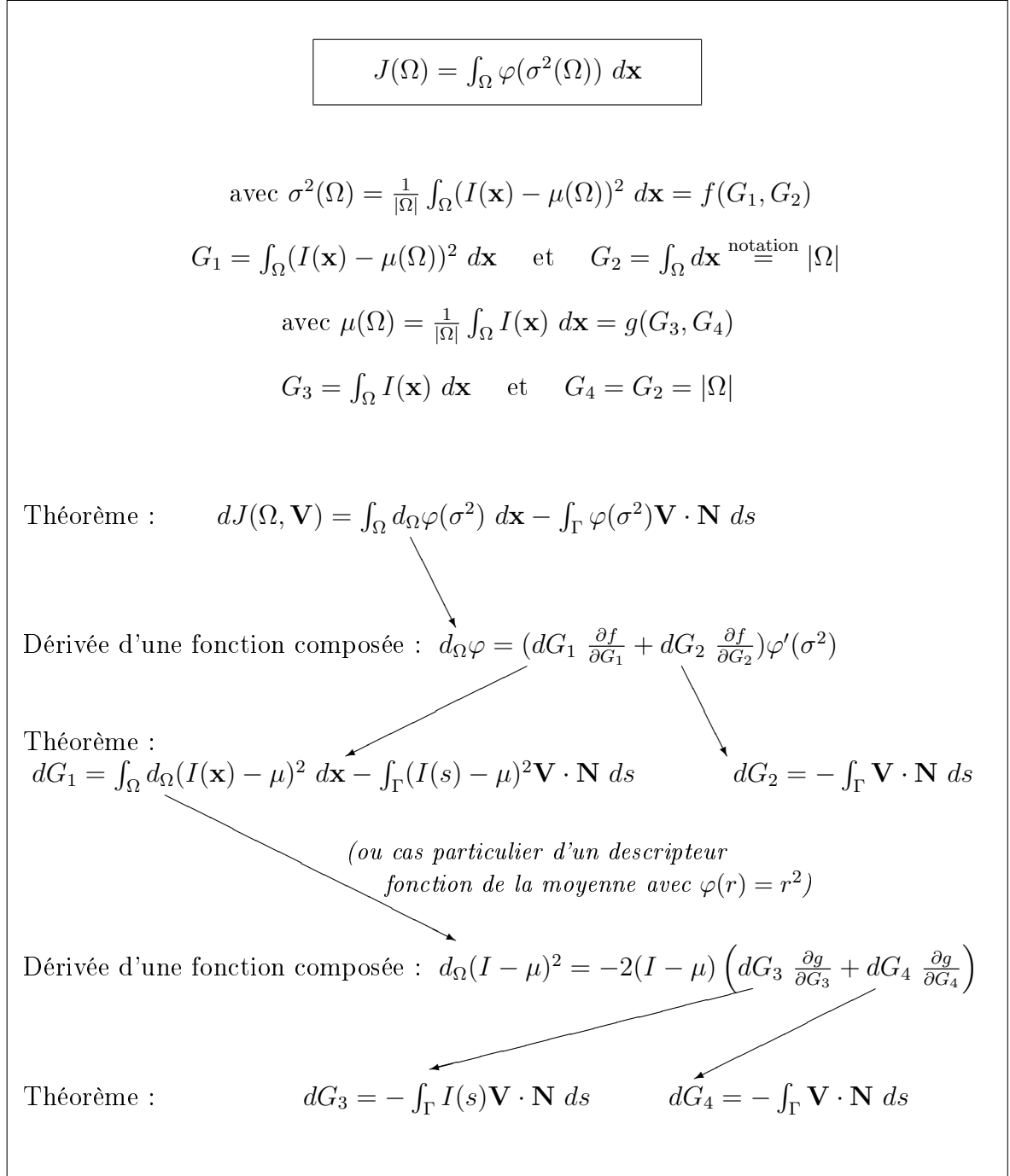


FIG. 3.3 – Étapes de dérivation d'un terme de région dont le descripteur est une fonction de la variance de la région.

Nous pourrions également utiliser le théorème 3.1 comme présenté dans la figure 3.3.

$$G_1(\Omega) = \int_{\Omega} (I(\mathbf{x}) - \mu(\Omega))^2 d\mathbf{x} \quad (3.46)$$

D'après les résultats de l'équation (3.37)

$$dG_1(\Omega, \mathbf{V}) = - \int_{\Gamma} (I(s) - \mu(\Omega))^2 \mathbf{V} \cdot \mathbf{N} ds \quad (3.47)$$

Pour calculer la dérivée eulérienne de

$$G_2(\Omega) = \int_{\Omega} d\mathbf{x} \quad (3.48)$$

nous utilisons une fois de plus le théorème 3.1 :

$$dG_2(\Omega, \mathbf{V}) = - \int_{\Gamma} \mathbf{V} \cdot \mathbf{N} ds \quad (3.49)$$

En remplaçant, dans l'équation (3.43), chacune des dérivées partielles de σ^2 par leur expression dans les équations (3.44) et (3.45), et chacune des dérivées eulériennes de G_1 et G_2 par leur expression dans les équations (3.47) et (3.49), on obtient l'expression de la dérivée eulérienne de σ^2 dans la direction \mathbf{V} :

$$\begin{aligned} d\sigma^2(\Omega, \mathbf{V}) &= - \int_{\Gamma} \frac{(I(s) - \mu(\Omega))^2}{|\Omega|} \mathbf{V} \cdot \mathbf{N} ds + \int_{\Gamma} \frac{\sigma^2(\Omega)}{|\Omega|} \mathbf{V} \cdot \mathbf{N} ds \\ &= - \int_{\Gamma} \frac{(I(s) - \mu(\Omega))^2 - \sigma^2(\Omega)}{|\Omega|} \mathbf{V} \cdot \mathbf{N} ds \end{aligned} \quad (3.50)$$

Finalement, on déduit des équations (3.40), (3.41) et (3.50) l'expression de la dérivée du critère :

$$\begin{aligned} dJ(\Omega, \mathbf{V}) &= - \int_{\Omega} \varphi'(\sigma^2(\Omega)) \int_{\Gamma} \frac{(I(s) - \mu(\Omega))^2 - \sigma^2(\Omega)}{|\Omega|} \mathbf{V} \cdot \mathbf{N} ds d\mathbf{x} \\ &\quad - \int_{\Gamma} \varphi(\sigma^2(\Omega)) \mathbf{V} \cdot \mathbf{N} ds \end{aligned} \quad (3.51)$$

soit :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} (\varphi(\sigma^2(\Omega)) + \varphi'(\sigma^2(\Omega)) [(I(s) - \mu(\Omega))^2 - \sigma^2(\Omega)]) \mathbf{V} \cdot \mathbf{N} ds \quad (3.52)$$

L'équation d'évolution associée est :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= [\varphi(\sigma^2(\Omega)) + \varphi'(\sigma^2(\Omega)) [(I(\mathbf{s}) - \mu(\Omega(\tau)))^2 - \sigma^2(\Omega)]] \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.53)$$

avec $\mathbf{s} = \Gamma(p, \tau)$.

3.4.2 Cas général : théorème

Cette section résume les différentes étapes du calcul de l'équation d'évolution issue d'un terme de région, et ce dans le cadre le plus général possible, c'est-à-dire lorsque le terme correspondant dépend doublement de la région. Le descripteur région est alors modélisé comme une combinaison de termes dépendants de la région. Nous présentons le théorème donnant l'expression de la dérivée du critère puis l'équation d'évolution du contour dans le cas général.

Nous vous invitons à consulter la thèse de Stéphanie Jehan-Besson [48] dans laquelle vous trouverez les démonstrations des théorèmes énoncés.

Tout d'abord, le descripteur région k est modélisé comme une combinaison linéaire de fonctionnelles dépendant elles-mêmes de la région :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x} = \int_{\Omega} k(\mathbf{x}, G_1(\Omega), G_2(\Omega), \dots, G_m(\Omega)) d\mathbf{x} \quad (3.54)$$

où m est un entier supérieur ou égal à 1 et où les fonctionnelles G_i dépendent aussi de la région :

$$G_i(\Omega) = \int_{\Omega} h_i(\mathbf{x}, \Omega) d\mathbf{x} \quad \text{pour } i = 1..m \quad (3.55)$$

Les fonctions h_i peuvent dépendre de la région et être exprimées comme une combinaison linéaire de p_i fonctions :

$$h_i(\mathbf{x}, \Omega) = h_i(\mathbf{x}, K_{i1}(\Omega), K_{i2}(\Omega), \dots, K_{ip_i}(\Omega)) \quad (3.56)$$

avec

$$K_{ij}(\Omega) = \int_{\Omega} l_{ij}(\mathbf{x}) d\mathbf{x} \quad \text{pour } i = 1..m, j = 1..p_i \quad (3.57)$$

Le procédé peut être réitéré, mais les descripteurs que nous utiliserons dans ce document ne nécessitent que deux niveaux de décomposition (ex : moyenne, variance).

Le théorème suivant donne l'expression de la dérivée du critère par rapport au paramètre d'évolution τ . Vous trouverez la démonstration de ce théorème dans la thèse de Stéphanie Jehan-Besson ([48] p. 44).

Théorème 3.3. Soit J le terme de région défini par

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) \, d\mathbf{x} = \int_{\Omega} k(\mathbf{x}, G_1(\Omega), G_2(\Omega), \dots, G_m(\Omega)) \, d\mathbf{x} \quad (3.58)$$

Sa dérivée eulérienne dans la direction \mathbf{V} est :

$$\begin{aligned} dJ(\Omega, \mathbf{V}) = - \int_{\partial\Omega} & \left(k(s, \Omega) + \sum_{i=1}^m (A_i h_i(s, \Omega)) \right. \\ & \left. + \sum_{i=1}^m A_i \sum_{j=1}^{p_i} (B_{ij} l_{ij}(s)) \right) (\mathbf{V} \cdot \mathbf{N}) \, ds \end{aligned} \quad (3.59)$$

avec pour $i = 1..m$:

$$A_i = \int_{\Omega} \frac{\partial k}{\partial G_i}(\mathbf{x}, G_1(\Omega), \dots, G_m(\Omega)) d\mathbf{x} \quad (3.60)$$

et pour $i = 1..m, j = 1..p_i$:

$$B_{ij} = \int_{\Omega} \frac{\partial h_i}{\partial K_{ij}}(\mathbf{x}, K_{i1}(\Omega), \dots, K_{ip_i}(\Omega)) d\mathbf{x} \quad (3.61)$$

et l'équation d'évolution du contour actif est :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= \left[k(\mathbf{s}, \Omega) + \sum_{i=1}^m (A_i h_i(\mathbf{s}, \Omega)) + \sum_{i=1}^m A_i \sum_{j=1}^{p_i} (B_{ij} l_{ij}(\mathbf{s})) \right] \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{aligned} \quad (3.62)$$

avec $\mathbf{s} = \Gamma(p, \tau)$. Le contour initial $\Gamma_0(p)$ est défini par l'utilisateur.

La vitesse d'évolution du contour actif \mathbf{V} est choisie telle que la dérivée du critère soit négative, ceci afin de s'assurer que le contour évolue vers le minimum du critère.

Chapitre 4

Mise en œuvre des contours actifs

La mise en œuvre d'une EDP gérant l'évolution d'un contour actif doit être stable et précise. Une simple approximation lagrangienne nécessite un pas de discrétisation très petit pour être stable. De plus, l'algorithme doit être capable de gérer les changements de topologie du contour : un contour doit pouvoir se séparer en deux objets, ou deux contours fusionner en un seul.

Les différentes mises en œuvre peuvent se distinguer en deux approches : explicites ou implicites. Les approches explicites définissent explicitement le contour à partir de paramètres. L'approche implicite des ensembles de niveau considère le contour comme le niveau zéro d'une fonction de dimension supérieure.

Ces deux approches seront succinctement décrites avec leurs avantages et leurs inconvénients.

4.1 Contours actifs paramétriques

Les approches paramétriques décrivent explicitement le contour à partir d'un certain nombre de paramètres.

Les *snakes* sont considérés comme le premier exemple de structure de contours actifs dont les propriétés intrinsèques sont prises en compte dans la minimisation de l'énergie. Les contours utilisés par Kass *et al.* [52] sont définis comme des courbes polygonales.

Les fonctions *B-splines* ont suscité un intérêt croissant pour leurs propriétés intrinsèques de régularité [17, 72].

Nous allons brièvement expliquer le principe d'une approche paramétrique avec l'exemple des *B-splines* (voir [71] pour de plus amples détails).

Commençons par définir la paramétrisation du contour.

Le contour actif est échantillonné par n points P_k .

La courbe $S_i(s)$, paramétrée par s entre les points P_i et P_{i+1} , est représentée par une *B-spline*.

La *B-spline* est choisie cubique car le calcul de la courbure nécessite de la dériver deux fois.

Une *B-spline* est définie par décomposition sur une base de fonctions $B_{S_i^j}$ déterminée par l'ordre de la *spline* (ici cubique, donc $j = 4$) :

$$\begin{aligned} S_i(s) &= \begin{pmatrix} x_i(s) \\ y_i(s) \end{pmatrix} \\ &= Q_{i-1}(s) * B_{S_{i-3}^4}(s) + Q_i(s) * B_{S_{i-2}^4}(s) \\ &\quad + Q_{i+1}(s) * B_{S_{i-1}^4}(s) + Q_{i+2}(s) * B_{S_i^4}(s) \end{aligned} \quad (4.1)$$

où les coefficients Q_i sont appelés les points de contrôle de la *spline*.

Les points de contrôle sont estimés à partir des points d'interpolation de la *spline*, qui sont les points d'échantillonnage du contour.

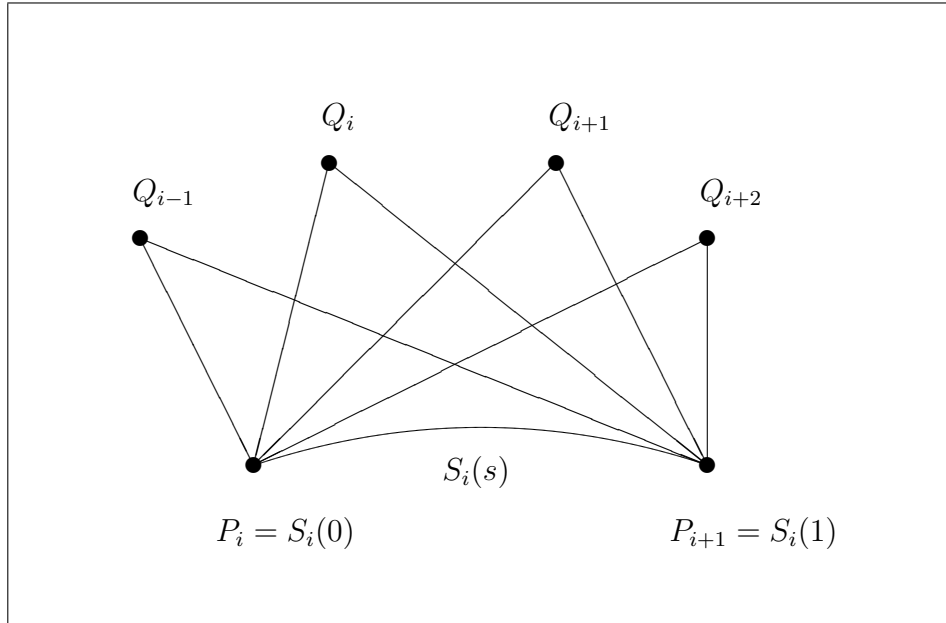


FIG. 4.1 – Mise en œuvre paramétrique : construction d'une B-spline cubique.

Si l'évolution du contour actif est donnée par :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = F(p, \tau) \mathbf{N}(p, \tau) \quad (4.2)$$

alors la force F sera appliquée à chaque point d'interpolation $P_i(\tau)$, dans la direction de la normale.

L'expression analytique de la *spline* (4.2) permet de calculer facilement certaines propriétés de la courbe nécessaires à l'évaluation de la force F , comme la courbure par exemple :

$$\kappa = \frac{x'' y' - x' y''}{(x'^2 + y'^2)^{\frac{3}{2}}} \quad (4.3)$$

Une nouvelle *spline* est définie à partir des nouveaux points d'interpolation $P_i(\tau + \delta\tau)$ et ceci jusqu'à convergence.

Les approches paramétriques ont pour principal avantage de réduire le nombre de points sur lesquels la force est calculée, et donc d'être rapides. Cependant, les changements de topologie doivent être gérés explicitement.

4.2 Méthode des ensembles de niveaux

La méthode des ensembles de niveaux [10, 58, 67] définit une paramétrisation implicite du contour.

On suppose que le contour actif est le niveau zéro d'une fonction de dimension supérieure $u : \mathbb{R}^+ \times \mathbb{R}^2 \rightarrow \mathbb{R}$, c'est-à-dire que

$$\mathbf{x} \in \Gamma(p, \tau) \iff u(\mathbf{x}, \tau) = 0 \quad (4.4)$$

La fonction u est classiquement choisie comme la fonction distance signée au contour $\Gamma(\tau)$, positive à l'extérieur du contour, et négative à l'intérieur.

Si le contour Γ évolue suivant l'équation :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = F(p, \tau) \mathbf{N}(p, \tau) \quad (4.5)$$

le niveau zéro de la fonction u évolue suivant l'équation :

$$\frac{\partial u(\Gamma(\tau, p), \tau)}{\partial \tau} = F(p, \tau) |\nabla u(\Gamma(\tau, p), \tau)| \quad (4.6)$$

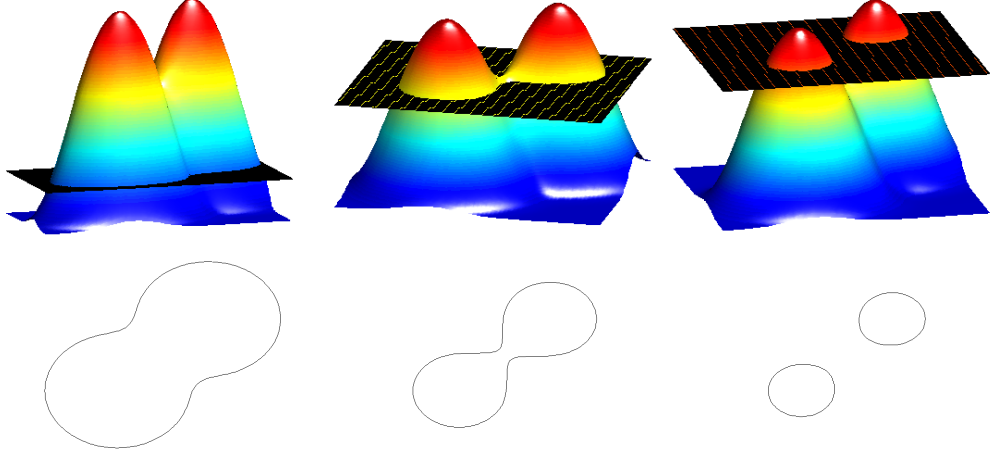


FIG. 4.2 – Méthode des ensembles de niveaux : gestion de la topologie. La première ligne d'images représente la carte des distances. Chaque carte est coupée en son niveau 0. Les contours correspondants constituent la deuxième ligne d'images.

L'équation (4.6) n'est valide que sur le niveau 0 de u , mais on peut étendre son expression à l'image si F y est définie :

$$\frac{\partial u(\mathbf{x}, \tau)}{\partial \tau} = F(\mathbf{x}, \tau) |\nabla u(\mathbf{x}, \tau)| \quad (4.7)$$

Cependant cette extension ne conserve pas la fonction de distance, et son gradient peut tendre vers l'infini.

Une solution à ce problème consiste à réinitialiser la fonction distance avec une EDP telle que $|\nabla u| = 1$ au cours de l'évolution :

$$\frac{\partial u(\mathbf{x}, \tau)}{\partial \tau} = \text{sign}(u(\mathbf{x}, \tau)) (1 - |\nabla u(\mathbf{x}, \tau)|) \quad (4.8)$$

où $\text{sign}(u(\mathbf{x}, \tau))$ représente le signe de $u(\mathbf{x}, \tau)$.

La dérivée partielle de u est discrétisée par le schéma numérique suivant :

$$\frac{\partial u(\mathbf{x}, \tau)}{\partial \tau} = \frac{u(\mathbf{x}, \tau + \delta\tau) - u(\mathbf{x}, \tau)}{\delta\tau} \quad (4.9)$$

où $\delta\tau$ représente le pas de discrétisation temporel.

A partir des équations (4.7) et (4.9) on déduit le schéma numérique de l'équation d'évolution :

$$u(\mathbf{x}, \tau + \delta\tau) = u(\mathbf{x}, \tau) + \delta\tau F(\mathbf{x}, \tau) |\nabla u(\mathbf{x}, \tau)| \quad (4.10)$$

L'avantage principal de la méthode des ensembles de niveaux réside dans la gestion implicite des changements de topologie (voir figure 4.2).

La représentation est Eulérienne, et la grille discrète définie lors de la discrétisation de l'équation d'évolution est fixe. Les schémas numériques sont stables et les résultats précis.

Les propriétés géométriques du contour actif telles que la courbure ou le vecteur normal peuvent être calculées à partir de la fonction u :

$$\mathbf{N} = -\frac{\nabla u}{|\nabla u|} \text{ et la courbure } \kappa = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) \quad (4.11)$$

La formulation peut être facilement étendue à des dimensions plus élevées.

L'inconvénient principal de la méthode des ensembles de niveaux est son coût calculatoire. Pour accélérer l'algorithme, on peut n'effectuer les calculs que sur une bande étroite $[1, 58]$ entourant le niveau 0 de u .

Deuxième partie

Segmentation avec a priori
géométrique sans contrainte
paramétrique

Contours actifs et a priori géométrique

5.1 Introduction et état de l'art

Dans de nombreux domaines d'applications des contours actifs, on dispose d'une information a priori sur la forme de l'objet à segmenter.

Cette information peut se traduire par un certain nombre de segmentations de l'objet d'intérêt, constituant un ensemble d'apprentissage.

En segmentation médicale, l'objet d'intérêt concerne des structures anatomiques semblables d'un individu à l'autre, et l'on souhaite définir une forme représentative de ces structures, un modèle. Les approches les plus courantes sont statistiques [8, 11, 14, 15, 16, 55, 83, 87]. Plus précisément, les méthodes présentées dans [8, 55, 83, 87] définissent un modèle paramétrique de forme par une analyse en composantes principales de l'ensemble des courbes. Suivant les travaux de Leventon *et al.* [55], dans [8, 83], l'analyse en composantes principales est effectuée sur un ensemble de représentations par cartes de distances des courbes d'apprentissage.

Rousson *et al.* [75] définissent aussi une représentation d'un ensemble de formes par leur carte de distance, dans le cadre de la segmentation d'images non médicales.

Charpiat *et al.* [14] utilisent les fonctions distance pour représenter un ensemble de formes dans un contexte statistique.

En segmentation de vidéos [35, 69], on considère que l'objet suivi se déforme peu d'une image à l'autre. L'information a priori est donnée par la segmentation de l'image précédente, recalée dans l'image courante.

L'information a priori peut simplement provenir d'un atlas anatomique,

d'une forme connue ou définie par un utilisateur. En segmentation [19, 30] ou pour des applications de déformation de contours ou shape warping [14, 35], elle se présente sous la forme d'un contour connu ou défini interactivement. En reconstruction de surface [91], il s'agira d'un ensemble de points ou d'un maillage.

La prise en compte de l'information a priori se formule comme la minimisation d'un critère traduisant une dissemblance entre le contour actif et le modèle de référence.

Dans certaines approches [30, 81, 83] le critère est paramétré par le choix de la représentation d'une courbe.

Les travaux pionniers de Staib *et al.* [81] décomposent le contour dans une base de Fourier. La minimisation du critère s'effectue par une descente de gradient sur l'espace des paramètres.

Plus récemment, Cremers *et al.* [22] représentent le contour actif par des B-Splines. Le descripteur de forme contient les coordonnées des points de contrôle de la spline. Leur critère minimise la distance de Mahalanobis entre les points de contrôle du contour actif et ceux de la forme de référence.

Foulonneau *et al.* [30] proposent une représentation paramétrique de la forme en projetant la fonction caractéristique de la forme sur la base des polynômes de Legendre. Leur critère minimise la distance quadratique entre les coefficients des contours actifs et de référence dans cet espace.

D'autres approches [15, 16, 55, 69, 75] utilisent une représentation implicite du contour actif. Ils définissent une contrainte paramétrique sur le contour actif par l'introduction d'une transformation entre le contour de référence et le contour actif. L'estimation des paramètres de la transformation est couplée au problème de segmentation.

Nous proposons un critère d'a priori de forme minimisant la distance Euclidienne entre le contour actif et un contour de référence. Notre contrainte ne fait pas l'hypothèse d'une transformation paramétrique entre le contour actif et le contour de référence. Le critère ne dépend pas d'une paramétrisation particulière du contour. L'équation d'évolution peut être implémentée implicitement ou paramétriquement. Nous avons choisi la méthode des ensembles de niveaux.

Dans ce chapitre, nous commencerons par définir les distances utilisées dans le critère, puis nous présenterons le critère d'a priori. Enfin, nous décrirons l'algorithme de contours actifs mis en oeuvre.

Cette contrainte géométrique sera étudiée seule pour des applications de déformation de contours, ou associée à des termes de région statistiques pour la segmentation et le suivi d'objets.

5.2 Définition de la contrainte géométrique non paramétrique

5.2.1 Distance à un contour : définition et propriétés directes

Soit une région Ω de l'image I , délimitée par son contour Γ , et Γ^{ref} un contour, qui sera dit de référence.

On définit la **fonction distance d d'un point \mathbf{x} au contour Γ^{ref}** comme suit :

$$\begin{aligned} d(\mathbf{x}, \Gamma^{ref}) &= \min_{\mathbf{y} \in \Gamma^{ref}} d(\mathbf{x}, \mathbf{y}) \\ &= \min_{\mathbf{y} \in \Gamma^{ref}} |\mathbf{x} - \mathbf{y}| \end{aligned} \quad (5.1)$$

La fonction $d(\mathbf{x}, \Gamma^{ref})$ est Lipschitz continue de constante 1. En effet :

$$|d(\mathbf{x}, \Gamma^{ref}) - d(\mathbf{y}, \Gamma^{ref})| \leq |\mathbf{x} - \mathbf{y}| \text{ pour tous } \mathbf{x}, \mathbf{y} \in I \quad (5.2)$$

ce qui implique suivant le théorème de Rademacher qu'elle est différentiable presque partout, sauf sur un ensemble négligeable.

Soit $\Pi_{\Gamma^{ref}}(\mathbf{x})$ l'ensemble des projections de \mathbf{x} sur Γ^{ref} par :

$$\Pi_{\Gamma^{ref}}(\mathbf{x}) = \{\mathbf{z} \in \Gamma^{ref} : |\mathbf{z} - \mathbf{x}| = d(\mathbf{x}, \Gamma^{ref})\}. \quad (5.3)$$

L'ensemble des points \mathbf{x} de I pour lesquels cette projection n'est pas unique est appelé le squelette de Γ^{ref} [26] (voir les figures 5.1 et 5.2).

Remarque 5.1. *La fonction $d(\mathbf{x}, \Gamma^{ref})$ n'est pas différentiable sur le squelette de Γ^{ref} ni sur Γ^{ref} lui-même.*

Enfin nous définissons la **distance signée à un contour Γ^{ref}** comme étant :

$$d(\mathbf{x}, \Gamma^{ref}) = \begin{cases} d(\mathbf{x}, \Gamma^{ref}), & \mathbf{x} \notin \Omega^{ref} \\ -d(\mathbf{x}, \Gamma^{ref}), & \mathbf{x} \in \Omega^{ref} \end{cases} \quad (5.4)$$

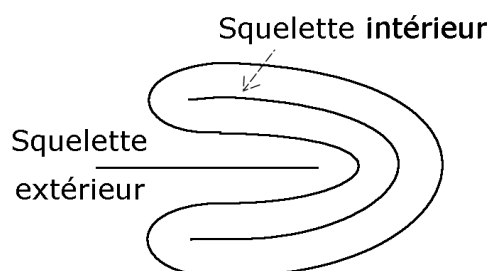


FIG. 5.1 – Squelette de Γ^{ref} : pour une forme concave, le squelette a une composante intérieure et au moins une composante extérieure

Si le point \mathbf{x} est à l'intérieur de Ω^{ref} , la région délimitée par le contour Γ^{ref} , alors sa distance au contour Γ^{ref} est définie négative. S'il est à l'extérieur de Ω^{ref} , sa distance au contour Γ^{ref} est positive (voir la figure 5.2).

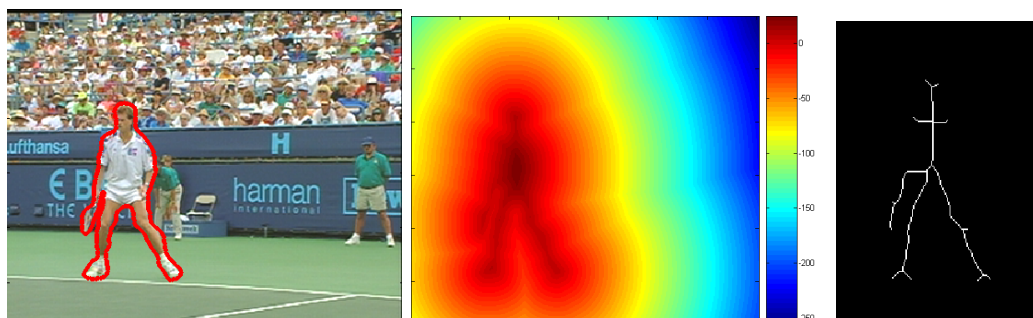


FIG. 5.2 – Illustrations sur une image réelle : de gauche à droite, un contour de référence (en rouge) détournant un objet, la carte de distance signée associée, et le squelette intérieur de l'objet.

5.2.2 Définition du critère d'a priori géométrique

L'information a priori se présente sous la forme d'un contour de référence, noté Γ^{ref} . Dans le cadre d'applications de type recalage, le contour de référence Γ^{ref} peut être extrait d'un atlas. Mais il peut également être défini interactivement par un utilisateur pour une application de segmentation, ou se déduire de la segmentation de l'image précédente pour des applications

de segmentation de vidéos (*tracking*).

Nous souhaitons contraindre le contour actif en minimisant la distance qui le sépare du contour de référence. Pour cela, nous définissons un critère J dont nous déduirons l'équation d'évolution du contour.

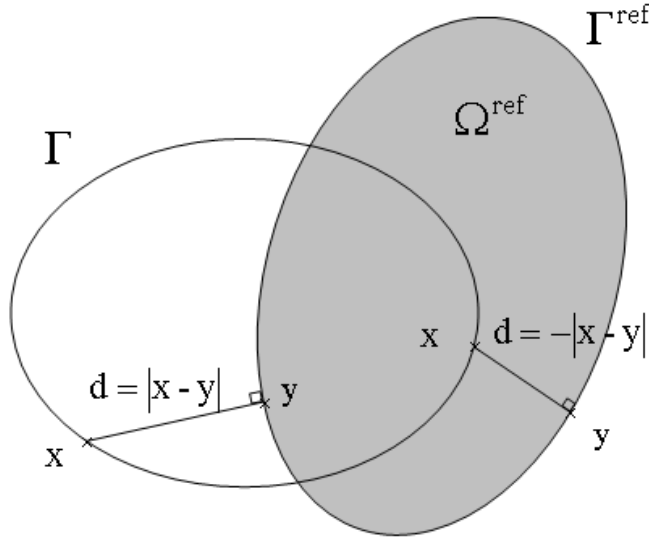


FIG. 5.3 – Fonction distance signée à un contour de référence

Le critère J s'écrit comme l'intégrale sur le contour Γ d'une fonction de la distance d entre un point $\mathbf{x} = \Gamma(s)$ du contour Γ et un contour de référence Γ^{ref} :

$$J(\Gamma) = \int_{\Gamma} \varphi(d(\mathbf{x}, \Gamma^{ref})) ds \quad (5.5)$$

où s représente l'abscisse curviligne du contour.

La fonction $d(\mathbf{x}, \Gamma^{ref})$ est définie par :

$$d(\mathbf{x}, \Gamma^{ref}) = \begin{cases} \min_{\mathbf{y} \in \Gamma^{ref}} (|\mathbf{x} - \mathbf{y}|) = |\mathbf{x} - \mathbf{y}(\mathbf{x})|, & \mathbf{x} \notin \Omega^{ref} \\ -\min_{\mathbf{y} \in \Gamma^{ref}} (|\mathbf{x} - \mathbf{y}|) = -|\mathbf{x} - \mathbf{y}(\mathbf{x})|, & \mathbf{x} \in \Omega^{ref} \end{cases} \quad (5.6)$$

où Ω^{ref} est la région intérieure délimitée par Γ^{ref} (voir la figure Fig. 5.3).

Classiquement, nous choisissons une fonction φ

- dérivable, car l'équation d'évolution du contour actif se déduit de la dérivée du critère,
- paire, afin de pondérer de façon identique les points à l'intérieur et ceux à l'extérieur du contour,
- et croissante sur \Re^+ pour pénaliser plus fortement les points éloignés du contour.

Ainsi, en cherchant à minimiser le critère J , nous pénalisons les grandes distances entre le contour Γ et le contour de référence Γ^{ref} . Le contour Γ tend donc à se rapprocher du contour de référence.

Cependant nous n'imposons **aucune contrainte paramétrique** sur le contour Γ . Certains travaux [16] utilisant des contraintes géométriques font l'hypothèse que le contour Γ se déduit du contour de référence à partir d'une transformation paramétrique. Nous ne faisons pas ce type d'hypothèse, et en ce sens, notre contrainte est plus libre.

5.3 Segmentation par contours actifs : une approche variationnelle

L'approche variationnelle consiste à déduire l'équation d'évolution du contour actif à partir de la dérivée du critère de segmentation.

Nous allons donc dériver le critère d'a priori géométrique. Deux méthodes de dérivation vous seront présentées en sections 5.3.2 et 5.3.3, pour aboutir à l'équation d'évolution en section 5.3.4.

5.3.1 Dérivation du critère d'a priori géométrique

Définissons dans un premier temps un espace de transformations

$$T_\tau : \Gamma \rightarrow \Gamma(\tau)$$

tel que

$$T_0(\Gamma) = \Gamma_0$$

où Γ_0 est le contour actif initial.

Nous introduisons ainsi un schéma dynamique qui permet d'écrire le critère comme une fonction de τ :

$$J(\Gamma(\tau)) = \int_{\Gamma(\tau)} \varphi(d(\mathbf{x}(\tau), \Gamma^{ref})) \, ds \quad (5.7)$$

avec $\mathbf{x}(\tau) = \Gamma(s, \tau)$.

Nous pouvons à présent dériver le critère par rapport au paramètre d'évolution τ du contour actif.

Théorème 5.1. *La dérivée eulérienne dans la direction \mathbf{V} d'un terme $J(\Gamma) = \int_{\Gamma} \varphi(d(\mathbf{x}, \Gamma^{ref})) ds$, où $\mathbf{x} = \Gamma(s)$, est donnée par l'expression :*

$$dJ(\Gamma, V) = \int_{\Gamma} (-\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) - \varphi(d)\kappa) \mathbf{V} \cdot \mathbf{N} ds \quad (5.8)$$

où \mathbf{N}^{ref} et \mathbf{N} représentent respectivement les vecteurs normalisés intérieurs du contour de référence et du contour actif, p paramétrise le contour actif, κ est la courbure du contour actif, et $d = d(\mathbf{x}, \Gamma^{ref})$ avec $\mathbf{x} = \Gamma(s)$.

L'équation d'évolution associée est donnée par :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = [\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) + \varphi(d)\kappa] \mathbf{N} \quad (5.9)$$

où $d = d(\mathbf{x}, \Gamma^{ref})$ avec $\mathbf{x} = \Gamma(p, \tau)$

Nous proposons deux démonstrations différentes de ce théorème : la première utilise le calcul des variations, la seconde les outils de dérivation de domaine, c'est-à-dire le gradient de forme.

5.3.2 Démonstration par le calcul des variations

Soit $J(\Gamma(\tau))$ défini par :

$$J(\Gamma(\tau)) = \int_{\Gamma(\tau)} \varphi(d(\mathbf{x}(\tau), \Gamma^{ref})) ds \quad (5.10)$$

où $d(\mathbf{x}(\tau), \Gamma^{ref})$ est la distance signée entre un point $\mathbf{x}(\tau) = \Gamma(s, \tau)$ du contour actif $\Gamma(\tau)$ et le contour de référence Γ^{ref} .

$$d(\mathbf{x}(\tau), \Gamma^{ref}) = \min_{\mathbf{y}_{ref} \in \Gamma^{ref}} (|\mathbf{x}(\tau) - \mathbf{y}_{ref}|) = |\mathbf{x}(\tau) - \mathbf{y}(\mathbf{x}(\tau))|, \quad \mathbf{x}(\tau) \in \Gamma(\tau) \quad (5.11)$$

Si le contour actif $\Gamma(\tau)$ est paramétré par p compris entre 0 et 1, le critère $J(\Gamma(\tau))$ s'écrit alors en fonction de $\mathbf{x}(p, \tau)$ et de $\mathbf{y}(p, \tau)$:

$$J(\Gamma(\tau)) = \int_0^1 \varphi(|\mathbf{x}(p, \tau) - \mathbf{y}(p, \tau)|) \left| \frac{\partial \mathbf{x}}{\partial p}(p, \tau) \right| dp \quad (5.12)$$

Pour simplifier les écritures des équations, nous remplacerons par la suite $\mathbf{x}(p, \tau)$ par \mathbf{x} et $\mathbf{y}(p, \tau)$ par \mathbf{y} , ainsi que $d(\mathbf{x}(\tau), \Gamma^{ref})$ par d .

Dérivons l'équation précédente (5.12) par rapport à τ . Nous obtenons :

$$\begin{aligned} dJ(\tau) &= \int_0^1 \left\langle \frac{\partial \mathbf{x}}{\partial \tau} - \frac{\partial \mathbf{y}}{\partial \tau}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \\ &+ \int_0^1 \varphi(d) \frac{\partial}{\partial \tau} \left(\left| \frac{\partial \mathbf{x}}{\partial p} \right| \right) dp \end{aligned} \quad (5.13)$$

où \langle , \rangle représente le produit scalaire.

Posons :

$$A = \int_0^1 \varphi(d) \frac{\partial}{\partial \tau} \left(\left| \frac{\partial \mathbf{x}}{\partial p} \right| \right) dp \quad (5.14)$$

En développant la dérivée partielle par rapport à τ , on obtient :

$$A = \int_0^1 \left\langle \varphi(d) \frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|}, \frac{\partial}{\partial p} \left(\frac{\partial \mathbf{x}}{\partial \tau} \right) \right\rangle dp \quad (5.15)$$

En intégrant par parties par rapport à p , on obtient :

$$A = - \int_0^1 \left\langle \frac{\partial}{\partial p} \left(\varphi(d) \frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} \right), \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle dp \quad (5.16)$$

Développons la dérivée $\frac{\partial}{\partial p} \left(\varphi(d) \frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} \right)$, A s'écrit alors :

$$A = - \int_0^1 \left\langle \frac{\partial}{\partial p} (\varphi(d)) \frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} + \varphi(d) \frac{\partial}{\partial p} \left(\frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} \right), \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle dp \quad (5.17)$$

Soit encore, en détaillant le terme $\frac{\partial}{\partial p} (\varphi(d))$:

$$A = - \int_0^1 \left\langle \left\langle \frac{\partial \mathbf{x}}{\partial p} - \frac{\partial \mathbf{y}}{\partial p}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} + \varphi(d) \frac{\partial}{\partial p} \left(\frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} \right), \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle dp \quad (5.18)$$

Notons \mathbf{N} la normale **intérieure** et \mathbf{T} la tangente de $\Gamma(\tau)$ (cf. fig. 5.4 page

54).

D'après les formules de Frenet, leurs expressions sont :

$$\mathbf{T} = \frac{\frac{\partial \mathbf{x}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} \text{ et } \kappa \mathbf{N} = \frac{1}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|} \frac{\partial \mathbf{T}}{\partial p} \quad (5.19)$$

où κ représente la courbure de $\Gamma(\tau)$.

Introduisons \mathbf{T} et $\kappa \mathbf{N}$ dans l'équation (5.18) :

$$A = - \int_0^1 \left\langle \left\langle \mathbf{T} - \frac{\frac{\partial \mathbf{y}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \mathbf{T} + \varphi(d) \kappa \mathbf{N}, \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \quad (5.20)$$

Reprenons l'expression entière de la dérivée du critère :

$$dJ(\tau) = \int_0^1 \left\langle \frac{\partial \mathbf{x}}{\partial \tau} - \frac{\partial \mathbf{y}}{\partial \tau}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp + A \quad (5.21)$$

En remplaçant, dans l'équation précédente, A par l'équation (5.20) nous obtenons :

$$\begin{aligned} dJ(\tau) &= \int_0^1 \left\langle \frac{\partial \mathbf{x}}{\partial \tau} - \frac{\partial \mathbf{y}}{\partial \tau}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \\ &\quad - \int_0^1 \left\langle \left\langle \mathbf{T} - \frac{\frac{\partial \mathbf{y}}{\partial p}}{\left| \frac{\partial \mathbf{x}}{\partial p} \right|}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \mathbf{T}, \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \\ &\quad - \int_0^1 \left\langle \varphi(d) \kappa \mathbf{N}, \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \end{aligned} \quad (5.22)$$

Lemme 5.1. *Les produits scalaires $\left\langle \frac{\partial \mathbf{y}}{\partial \tau}, \mathbf{x} - \mathbf{y} \right\rangle$ et $\left\langle \frac{\partial \mathbf{y}}{\partial p}, \mathbf{x} - \mathbf{y} \right\rangle$ sont nuls.*

Démonstration :

Pour tout $\mathbf{x} \in \Gamma$, par définition, \mathbf{y} appartient Γ^{ref} .

Comme Γ^{ref} est fixé, les dérivées partielles $\frac{\partial \mathbf{y}}{\partial \tau}$ et $\frac{\partial \mathbf{y}}{\partial p}$ sont nécessairement colinéaires à \mathbf{T}^{ref} .

Or, par définition, $\mathbf{x} - \mathbf{y} = \alpha \mathbf{N}^{ref}$, où α est une constante et \mathbf{N}^{ref} la normale

unitaire intérieure au contour de référence.

On en déduit que les produits scalaires $\langle \frac{\partial \mathbf{y}}{\partial \tau}, \mathbf{x} - \mathbf{y} \rangle$ et $\langle \frac{\partial \mathbf{y}}{\partial p}, \mathbf{x} - \mathbf{y} \rangle$ sont nuls.

□

Reprenons l'équation (5.22) et simplifions-la à l'aide du lemme précédent 5.1 :

$$\begin{aligned}
 dJ(\tau) &= \int_0^1 \left\langle \frac{\partial \mathbf{x}}{\partial \tau}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \\
 &\quad - \int_0^1 \left\langle \mathbf{T}, \frac{\mathbf{x} - \mathbf{y}}{d} \right\rangle \varphi'(d) \mathbf{T}, \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \\
 &\quad - \int_0^1 \left\langle \varphi(d) \kappa \mathbf{N}, \frac{\partial \mathbf{x}}{\partial \tau} \right\rangle \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp
 \end{aligned} \tag{5.23}$$

Notons \mathbf{N}^{ref} la normale unitaire **intérieure** du contour de référence Γ^{ref} (cf. fig. 5.4).

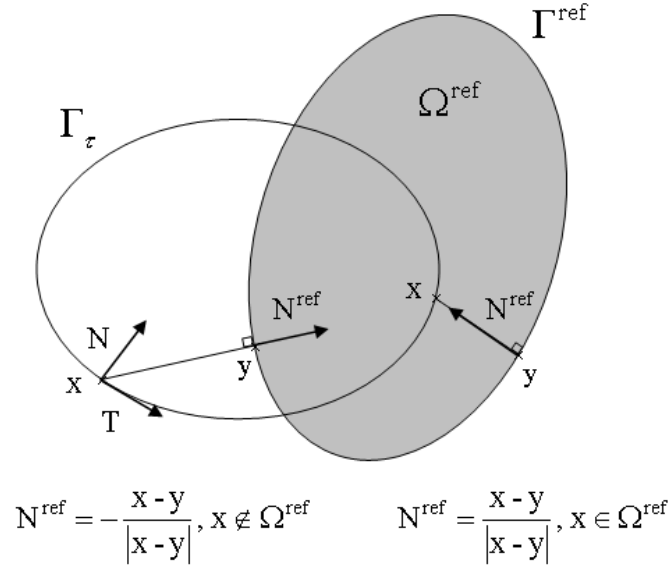


FIG. 5.4 – Normales et tangentes aux contours actifs et de référence.

Dans le cas où le point $\mathbf{x} \in \Gamma$ est extérieur à la région Ω^{ref} :

$$\mathbf{N}^{ref} = -\frac{\mathbf{x} - \mathbf{y}}{|\mathbf{x} - \mathbf{y}|}$$

Dans le cas où le point $\mathbf{x} \in \Gamma$ est intérieur à la région Ω^{ref} , la normale intérieure au contour de référence vaut :

$$\mathbf{N}^{ref} = \frac{\mathbf{x} - \mathbf{y}}{|\mathbf{x} - \mathbf{y}|}$$

On peut donc exprimer la normale intérieure au contour de référence pour tout \mathbf{x} par :

$$\mathbf{N}^{ref} = -\frac{\mathbf{x} - \mathbf{y}}{d} \quad (5.24)$$

Factorisons l'équation (5.23) puis introduisons \mathbf{N}^{ref} :

$$dJ(\tau) = \int_0^1 \left\langle \frac{\partial \mathbf{x}}{\partial \tau}, -\varphi'(d)\mathbf{N}^{ref} + \langle \mathbf{T}, \mathbf{N}^{ref} \rangle \varphi'(d)\mathbf{T} - \varphi(d)\kappa\mathbf{N} \right\rangle \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \quad (5.25)$$

Le vecteur \mathbf{N}^{ref} est alors décomposé sur la base (\mathbf{N}, \mathbf{T}) :

$$\mathbf{N}^{ref} = \mathbf{N}^{ref} \cdot \mathbf{N} \mathbf{N} + \mathbf{N}^{ref} \cdot \mathbf{T} \mathbf{T} \quad (5.26)$$

ce qui nous permet de remplacer, dans l'équation de la dérivée du critère, le produit $\varphi'(d)\mathbf{N}^{ref}$ par l'expression

$$\varphi'(d) \mathbf{N}^{ref} \cdot \mathbf{N} \mathbf{N} + \varphi'(d) \mathbf{N}^{ref} \cdot \mathbf{T} \mathbf{T} \quad (5.27)$$

Dans l'équation (5.25) la composante en \mathbf{T} s'annule et la dérivée du critère s'écrit :

$$dJ(\tau) = \int_0^1 \left\langle \frac{\partial \mathbf{x}}{\partial \tau}, (-\varphi'(d) \mathbf{N}^{ref} \cdot \mathbf{N} - \varphi(d) \kappa) \mathbf{N} \right\rangle \left| \frac{\partial \mathbf{x}}{\partial p} \right| dp \quad (5.28)$$

soit finalement :

$$dJ(\tau) = \int_{\Gamma(\tau)} (-\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) - \varphi(d) \kappa) \mathbf{V} \cdot \mathbf{N} ds \quad (5.29)$$

5.3.3 Démonstration par les gradients de forme

Soit un terme de contour de forme générale :

$$J(\Gamma) = \int_{\Gamma} k(s) ds \quad (5.30)$$

où $k(s)$ est un descripteur du contour Γ , contour paramétré par son abscisse curviligne.

La dérivée eulérienne du critère (5.30) dans la direction \mathbf{V} , telle que présentée par [10, 51, 38], vaut :

$$dJ(\Gamma, \mathbf{V}) = \int_{\Gamma} (dk(s, \mathbf{N}) - k(s) \kappa) \mathbf{V} \cdot \mathbf{N} ds \quad (5.31)$$

où \mathbf{N} est la normale unitaire intérieure au contour, et κ la courbure du contour.

La dérivée directionnelle $dk(s, \mathbf{N})$ peut s'écrire comme le produit scalaire du gradient de k et de la normale \mathbf{N} , soit :

$$dk(s, \mathbf{N}) = \nabla k \cdot \mathbf{N} \quad (5.32)$$

Dans notre cas, le descripteur du contour est une fonction de la distance au contour de référence :

$$k(s) = \varphi(d(\mathbf{x}, \Gamma^{ref})) \quad (5.33)$$

avec $\mathbf{x} = \Gamma(s)$.

La dérivée directionnelle du descripteur, vu comme une fonction composée, vaut alors :

$$dk(s, \mathbf{N}) = dd(\mathbf{x}, \mathbf{N}) \varphi'(d) \quad (5.34)$$

$$= \nabla d \cdot \mathbf{N} \varphi'(d) \quad (5.35)$$

or $\nabla d = \frac{\mathbf{x} - \mathbf{y}}{d} = -\mathbf{N}^{ref}$ (cf eq. (5.24)), on en déduit donc :

$$dk(s, \mathbf{N}) = -\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) \quad (5.36)$$

Finalement, la dérivée du critère (5.30) se déduit des équations (5.31) et (5.36) :

$$dJ(\Gamma, V) = \int_{\Gamma} (-\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) - \varphi(d) \kappa) \mathbf{V} \cdot \mathbf{N} ds \quad (5.37)$$

5.3.4 Équation d'évolution du contour actif

L'équation d'évolution du contour actif se déduit à partir de la dérivée du critère.

On cherche à faire évoluer le contour de sorte qu'il minimise le critère.

En choisissant l'équation d'évolution suivante :

$$\begin{aligned}\frac{\partial \Gamma(p, \tau)}{\partial \tau} &= [\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) + \varphi(d) \kappa] \mathbf{N} \\ \Gamma(0, \tau) &= \Gamma_0(\tau)\end{aligned}\tag{5.38}$$

avec $d = d(\mathbf{x}, \Gamma_{ref})$ et $\mathbf{x} = \Gamma(p, \tau)$,

on s'assure que la dérivée du critère $J(\Gamma(\tau))$ est négative, et donc que l'évolution du contour $\Gamma(\tau)$ fera décroître la valeur du critère $J(\Gamma(\tau))$.

Remarque 5.2. *La vitesse du contour actif doit s'annuler sur le contour de référence, c'est-à-dire pour $d = 0$. Ceci implique que la fonction φ doit satisfaire les contraintes suivantes :*

$$\begin{cases} \varphi(0) = 0 \\ \varphi'(0) = 0 \end{cases}\tag{5.39}$$

Remarque 5.3. *L'utilisation de la fonction φ permet d'éviter le calcul du gradient de la distance $d(\mathbf{x}, \Gamma^{ref})$, avec $\mathbf{x} = \Gamma(s)$, distance qui, rappelons-le, n'est pas différentiable sur le squelette de Γ^{ref} .*

Remarque 5.4. *L'équation (5.38) s'écrit comme la somme de deux termes : un terme hyperbolique $-\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d)\mathbf{N}$ et un terme parabolique $\varphi(d)\kappa\mathbf{N}$. Afin d'assurer la stabilité du schéma numérique, la discrétisation du terme hyperbolique doit satisfaire la condition de Courant-Friedrich-Lewy (CFL) [18]*

$$|\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d)\mathbf{N}| \frac{\delta \tau}{\delta \mathbf{x}} \leq 1\tag{5.40}$$

Une condition suffisante pour assurer la stabilité consiste à choisir φ de sorte que sa dérivée soit bornée et $\delta \tau$, le pas de discrétisation, tel que :

$$\delta \tau \leq \frac{\delta \mathbf{x}}{\max(\varphi')}\tag{5.41}$$

Chapitre 6

Applications

L'objectif de ce chapitre est de présenter différentes utilisations du critère de contrainte géométrique.

Tout d'abord, le critère sera étudié seul pour des applications de type déformation de contour ou *shape warping*. Ensuite, nous montrerons comment l'ajout d'une contrainte géométrique peut améliorer les résultats obtenus, en segmentation d'images et de vidéos, par rapport à des méthodes utilisant uniquement des descripteurs statistiques.

6.1 Déformation de contour

Les premières applications ne prennent en compte que les informations de distance à un contour de référence Γ^{ref} , défini arbitrairement par un utilisateur, ou provenant de segmentations préalables.

Cette distance est définie comme suit (voir section 5.2) :

$$d(\mathbf{x}, \Gamma^{ref}) = \begin{cases} \min_{\mathbf{y} \in \Gamma^{ref}} (|\mathbf{x} - \mathbf{y}|) = |\mathbf{x} - \mathbf{y}(\mathbf{x})|, & \mathbf{x} \notin \Omega^{ref} \\ -\min_{\mathbf{y} \in \Gamma^{ref}} (|\mathbf{x} - \mathbf{y}|) = -|\mathbf{x} - \mathbf{y}(\mathbf{x})|, & \mathbf{x} \in \Omega^{ref} \end{cases} \quad (6.1)$$

où Ω^{ref} est la région intérieure délimitée par Γ^{ref} .

Le critère s'écrit de manière générale :

$$J(\Gamma) = \int_{\Gamma} \varphi(d(\mathbf{x}, \Gamma^{ref})) \, ds \quad (6.2)$$

où $\mathbf{x} = \Gamma(s)$, et s est l'abscisse curviligne de Γ .

L'équation d'évolution du contour actif reprend donc directement celle calculée à la section 5.3.4 :

$$\begin{aligned}\frac{\partial \Gamma(p, \tau)}{\partial \tau} &= [\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) + \varphi(d)\kappa] \mathbf{N} \\ \Gamma(0, \tau) &= \Gamma_0(\tau)\end{aligned}\tag{6.3}$$

où $d = d(\mathbf{x}, \Gamma^{ref})$, $\mathbf{x} = \Gamma(p, \tau)$, \mathbf{N}^{ref} et \mathbf{N} sont les normales unitaires intérieures respectivement du contour de référence et du contour actif, et où κ est la courbure du contour actif.

La déformation de contour, ou *shape warping*, consiste à faire évoluer un contour à partir d'une forme A vers une autre forme B de façon continue. Dans notre cas, les deux formes en question sont soit définies manuellement, soit issues d'une segmentation préalable.

Le contour déformé est un contour actif, initialisé à la première forme A , et évoluant suivant l'équation (6.3) vers la forme B , considérée comme le contour de référence.

Nous allons d'abord nous pencher sur le choix de la fonction φ de sorte qu'elle satisfasse les conditions énoncées dans la section 5.2.2 et dans les remarques 5.2 et 5.4.

Ensuite nous présenterons le processus de déformation d'une ellipse verticale vers une ellipse horizontale et celles de la déformation d'un visage vers un autre.

6.1.1 Choix de la fonction de pondération de la distance

Rappelons les hypothèses contraignant le choix de la fonction de pondération φ .

La fonction φ doit être :

- dérivable, car l'équation d'évolution du contour actif se déduit de la dérivée du critère,
- paire, afin de pondérer à l'identique les points à l'intérieur et ceux à l'extérieur du contour,
- croissante sur \mathbb{R}^+ pour pénaliser plus fortement les points éloignés du contour,
- nulle et de dérivée nulle en zéro, c'est-à-dire pour tout point appartenant au contour de référence (voir Remarque 5.2),
- de dérivée bornée, condition suffisante pour assurer la stabilité du schéma numérique (voir Remarque 5.4).

La fonction φ choisie est : $\varphi(d) = \sqrt{1 + d^2} - 1$ [13].

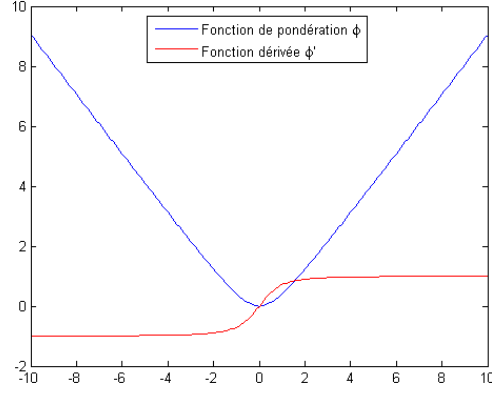


FIG. 6.1 – Fonction de pondération de la distance.

Le critère (6.2) s'écrit :

$$J(\Gamma) = \int_{\Gamma} \sqrt{1 + d^2(s, \Gamma^{ref})} - 1 \quad ds \quad (6.4)$$

et l'équation d'évolution correspondante est :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = [\mathbf{N}^{ref} \cdot \mathbf{N} \frac{d}{\sqrt{1 + d^2}} + (\sqrt{1 + d^2} - 1) \kappa] \mathbf{N} \quad \Gamma(0, \tau) = \Gamma_0(\tau) \quad (6.5)$$

Comme $\varphi'(d) = \frac{d}{\sqrt{1 + d^2}}$, le contour s'élargit ou se rétrécit suivant s'il est respectivement à l'intérieur ($d < 0$) ou à l'extérieur ($d > 0$) du contour de référence.

De plus, $\varphi'(d)$ est majoré par 1, ce qui est suffisant pour satisfaire la condition CFL si $\delta\tau$ est tel que : $\delta\tau \leq \frac{\delta\mathbf{x}}{\max(\varphi')}$. La stabilité du schéma numérique est donc bien assurée.

6.1.2 Mise en œuvre

L'algorithme mis en oeuvre pour cette application est le suivant :

- Initialisation des données :

- *Contour de référence* :
Le contour de référence Γ^{ref} peut être initialisé interactivement par un utilisateur mais peut également être extrait d'un atlas ou d'une segmentation préalable.
- *Contour actif* :
Le contour actif initial est défini par l'utilisateur.
- *Carte des distances* :
La carte des distances au contour de référence Γ^{ref} est calculée en utilisant un algorithme rapide présenté dans [86] (On aurait pu aussi utiliser l'équation (6.9)). La carte des distances est définie négative pour les pixels appartenant à l'objet Ω^{ref} , et positive ailleurs.
- *Évolution du contour actif* :
Pour chaque itération et ce jusqu'à convergence nous calculons :
 - *La vitesse F* : suivant l'équation

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = [\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) + \varphi(d)\kappa] \mathbf{N} = F(p, \tau)\mathbf{N} \quad (6.6)$$

- *Mise à jour de la fonction u* : suivant l'équation

$$\frac{\partial u(\mathbf{x}, \tau)}{\partial \tau} = F(\mathbf{x}, \tau)|\nabla u(\mathbf{x}, \tau)| \quad (6.7)$$

qui devient, après discrétisation :

$$u(\mathbf{x}, \tau + \delta\tau) = u(\mathbf{x}, \tau) + \delta\tau F(\mathbf{x}, \tau) |\nabla u(\mathbf{x}, \tau)| \quad (6.8)$$

Reinitialisation de u : u est régulièrement réinitialisée afin de rester une carte des distances, suivant l'équation (voir [82]) :

$$\frac{\partial u(\mathbf{x}, \tau)}{\partial \tau} = \text{sign}(u(\mathbf{x}, \tau))(1 - |\nabla u(\mathbf{x}, \tau)|) \quad (6.9)$$

La gestion d'éventuels changements de topologie est intrinsèque à la méthode des ensembles de niveaux. Vous trouverez plus de détails sur la méthode des ensembles de niveaux en section 4.2 page 39.

6.1.3 Résultats expérimentaux

Nous nous intéressons tout d'abord à un exemple simple de déformation. Pour cela nous initialisons un contour actif par une ellipse (rose sur la figure 6.2) et prenons comme contour de référence la limite, elliptique, entre un objet (en blanc) et le fond de l'image (en noir).

La figure 6.2 présente différentes itérations de l'algorithme de convergence

du contour actif vers le contour de référence. L'algorithme converge en 350 itérations environ pour une image de taille $150 * 150$ pixels et un pas de discrétisation $\delta\tau = 0.05$.

On peut voir que le contour actif se déforme continûment du contour initial au contour final.

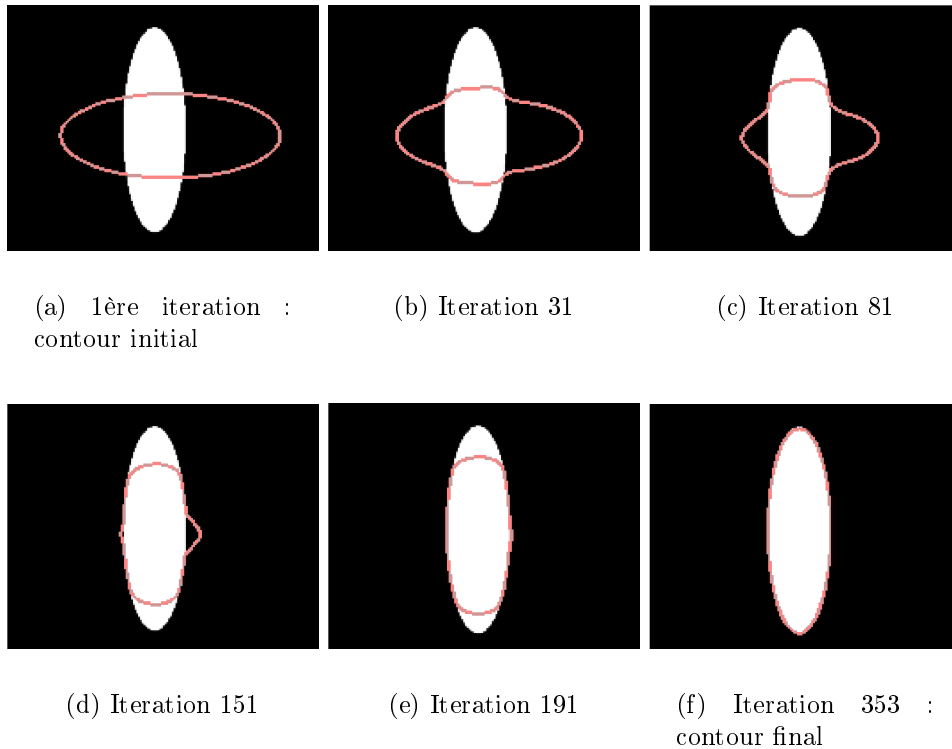


FIG. 6.2 – Déformations d'un contour actif évoluant d'une ellipse (en rose) vers une autre (définie par la limite entre l'objet blanc et le fond noir).

Ensuite nous avons testé le *shape warping* entre deux images réelles des séquences Foreman et Erik.

Le contour actif est initialisé autour du visage de Foreman, comme présenté dans la figure 6.3.a. Puis le contour de référence est défini comme délimitant le visage de Erik tel que dans la figure 6.3.b.

La figure 6.4 présente l'évolution du contour actif du contour initial vers le contour de référence. L'algorithme converge en 700 itérations environ, pour

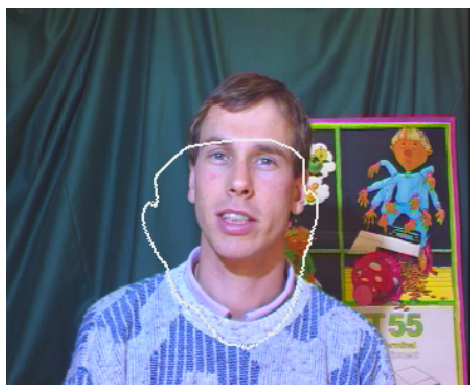


(a) Contour initial autour du visage de Foreman

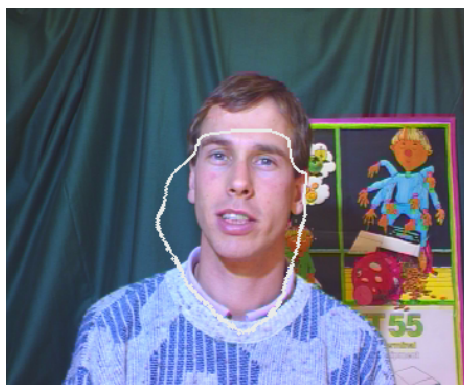
(b) Contour de référence autour du visage de Erik

FIG. 6.3 – Déformation d'un contour d'un visage à un autre : définition des contours initial et de référence

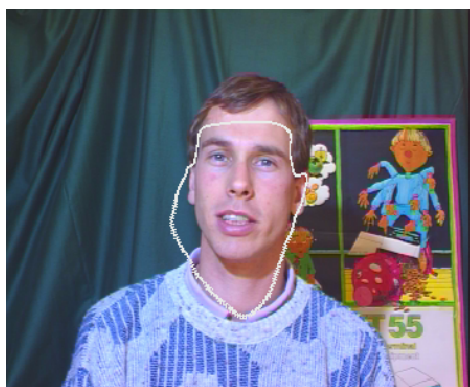
un pas de discrétisation $\delta\tau = 0.07$, et une image de taille $288 * 352$ pixels. Ces résultats illustrent le fait qu'il n'y a aucune contrainte paramétrique de forme sur le contour actif. Le contour se rapproche du contour de référence en se déformant librement, c'est-à-dire qu'on n'impose pas la contrainte d'une transformation géométrique (affine par exemple) permettant de passer du contour actif au contour de référence.



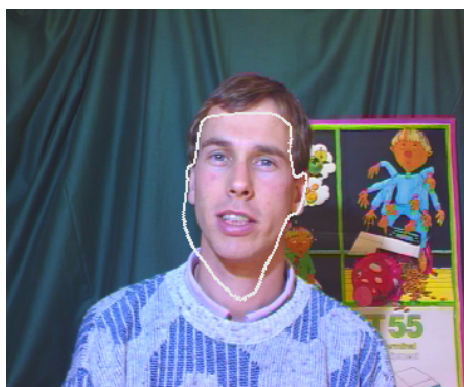
(a) 1ère itération



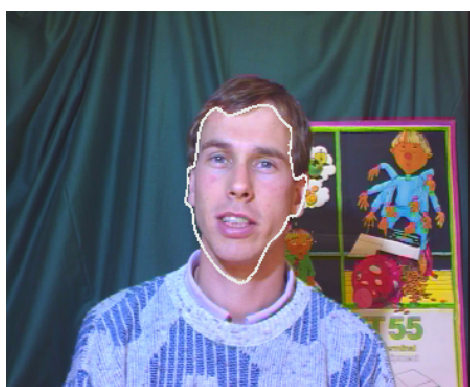
(b) Itération 100



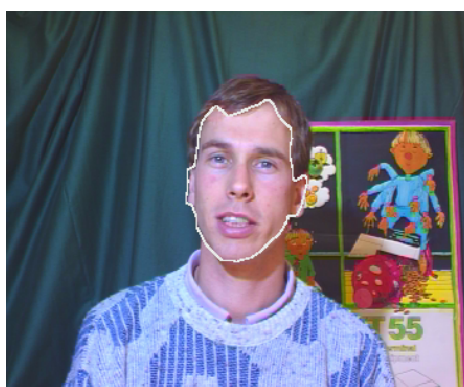
(c) Itération 200



(d) Itération 300



(e) Itération 400



(f) Itération finale 720

FIG. 6.4 – Déformation d'un contour d'un visage à un autre : évolution du contour actif

6.2 Segmentation interactive d'images

Pour des applications de segmentation d'objets, nous combinerons le terme de distance à un contour de référence, traduisant une connaissance a priori du contour, avec des termes régions statistiques et un terme de régularisation classique. Nous comparerons notre méthode à une compétition de régions utilisant le même critère statistique afin de montrer l'intérêt du terme d'a priori.

Dans cette section nous présenterons d'abord le critère de compétition de régions, puis nous définirons notre critère conjuguant terme a priori et terme statistique, enfin nous montrerons des résultats comparatifs des deux méthodes.

6.2.1 Critère de compétition de régions

Soit I une image composée d'une région d'intérêt Ω^{in} , appelée objet, et d'un fond Ω^{out} constituant une partition de l'image. Le contour Γ est la frontière commune à Ω^{in} et Ω^{out} .

La méthode de compétition de régions consiste à définir un critère qui pondère les informations extraites de chacune des deux régions objet et fond, les effets de chaque terme de région devant s'équilibrer. Un terme de contour peut s'ajouter pour compléter le critère.

Un tel critère s'écrit généralement :

$$\begin{aligned} J(\Gamma) = & \int_{\Omega^{in}} k^{in}(\mathbf{x}, \Omega^{in}) d\mathbf{x} \\ & + \lambda \int_{\Omega^{out}} k^{out}(\mathbf{x}, \Omega^{out}) d\mathbf{x} \\ & + \epsilon \int_{\Gamma} k^b(s) ds \end{aligned} \quad (6.10)$$

où $k^{in}(\mathbf{x}, \Omega^{in})$, $k^{out}(\mathbf{x}, \Omega^{out})$ et $k^b(s)$ sont respectivement le descripteur de l'objet, du fond et du contour. Les constantes λ et ϵ pondèrent l'influence de chacun des termes du critère.

Pour une application de segmentation de régions homogènes, nous avons choisi des descripteurs statistiques de région.

Les descripteurs contour et région sont :

$$\begin{cases} k^{out} &= \psi(S(\Omega^{out})) \\ k^{in} &= \psi(S(\Omega^{in})) \\ k^b &= 1 \end{cases} \quad (6.11)$$

où

- $S(\Omega) = \sigma_1^2(\Omega) + \sigma_2^2(\Omega) + \sigma_3^2(\Omega)$ est la somme des variances de chaque composante couleur de I sur la région Ω ,
- $\sigma_i^2(\Omega)$ est la variance de I pour la composante couleur i sur la région $\Omega = \Omega^{in}$ ou Ω^{out} ,
- et ψ est une fonction de pondération différentiable.
Nous avons choisi $\psi = \log(1 + r^2)$ ($\psi' = \frac{2r}{1+r^2}$).

Notons :

- $I_i(\mathbf{x})$ la luminance ($i = 1$) ou l'une des deux chrominances ($i = 2, 3$) de l'image I au pixel \mathbf{x} ,
- et $\mu_i(\Omega)$ la valeur moyenne de I pour la composante couleur i sur la région $\Omega = \Omega^{in}$ ou Ω^{out} .

Le terme de contour minimise la longueur du contour, c'est donc un terme de régularisation.

Le terme de région décrivant l'objet minimise l'intégrale sur la région Ω^{in} d'une fonction de la variance. Un tel terme utilisé seul tendrait à minimiser la région d'intégration Ω^{in} et ainsi à sous-segmenter la région d'intérêt, voire à faire rétrécir Ω^{in} jusqu'à disparaître.

De même, si le terme de région décrivant le fond de l'image n'était associé à aucun autre, la région Ω^{out} aurait tendance à rétrécir. L'objet serait cette fois sur-segmenté.

C'est pourquoi ces deux termes sont mis en compétition, afin d'avoir le meilleur compromis entre leurs effets contradictoires mais complémentaires.

L'équation d'évolution se déduit des théorèmes généraux de la section 3.4.2 ou des résultats obtenus à la section 3.4.1 pour la dérivation d'un terme de région de descripteur une fonction de la variance. Dans cette dernière, nous montrons que la dérivée d'un terme de région tel que :

$$J(\Omega) = \int_{\Omega} \varphi(\sigma^2(\Omega)) \, d\mathbf{x} \quad (6.12)$$

où $\sigma^2(\Omega)$ représente la variance de la région Ω , a pour expression :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} (\varphi(\sigma^2(\Omega)) + \varphi'(\sigma^2(\Omega))[(I(s) - \mu(\Omega))^2 - \sigma^2(\Omega)]) \mathbf{V} \cdot \mathbf{N} \, ds \quad (6.13)$$

Dans notre exemple, on a :

$$\varphi(\sigma^2(\Omega)) = \psi \left(\sum_{i=1}^3 \sigma_i^2(\Omega) \right) \quad (6.14)$$

La dérivée d'un des termes de région vaut donc :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} \psi(S(\Omega)) + \psi'(S(\Omega)) \sum_{i=1}^3 [(I_i(s) - \mu_i(\Omega))^2 - S(\Omega)] \mathbf{V} \cdot \mathbf{N} \, ds \quad (6.15)$$

pour $\Omega = \Omega^{in}$ ou Ω^{out} .

La dérivée d'un terme de contour $J(\Gamma) = \int_{\Gamma} k(s) \, ds$ est, rappelons-le :

$$dJ(\Gamma, \mathbf{V}) = \int_{\Gamma} (\nabla k(s) \cdot \mathbf{N} - k(s) \kappa) \mathbf{V} \cdot \mathbf{N} \, ds \quad (6.16)$$

Dans notre cas $k(s) = 1$, la dérivée du terme de contour s'écrit donc :

$$dJ(\Gamma, \mathbf{V}) = - \int_{\Gamma} \kappa \mathbf{V} \cdot \mathbf{N} \, ds \quad (6.17)$$

Finalement, le contour actif vérifiant le critère (6.10) suit l'équation d'évolution suivante :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= \psi(S(\Omega^{in}(\tau))) - \lambda \psi(S(\Omega^{out}(\tau))) \\ &+ \psi'(S(\Omega^{in}(\tau))) \sum_{i=1}^3 [(I_i(\mathbf{x}) - \mu_i(\Omega^{in}(\tau)))^2 - S(\Omega^{in}(\tau))] \\ &- \lambda \psi'(S(\Omega^{out}(\tau))) \sum_{i=1}^3 [(I_i(\mathbf{x}) - \mu_i(\Omega^{out}(\tau)))^2 - S(\Omega^{out}(\tau))] \\ &+ \epsilon \kappa] \mathbf{N} \end{aligned} \quad (6.18)$$

avec $\mathbf{x} = \Gamma(p, \tau)$.

Dans le cadre d'une segmentation faisant l'hypothèse d'un objet homogène, le critère ne peut pas contenir qu'un seul terme de région, car alors la région en question diminuerait jusqu'à disparaître. On doit mettre en compétition les informations extraites des deux régions que sont l'objet et le fond de l'image, pour trouver un équilibre entre les deux.

6.2.2 Critère combinant a priori géométrique et descripteur de région

L'introduction d'un a priori, sous la forme d'un terme de contour fondé sur la distance à un contour de référence, nous dispense d'utiliser un terme de

région décrivant le fond de l'image. En effet, cet a priori entre en compétition avec les informations statistiques de l'objet, et les informations statistiques du fond de l'image ne sont plus nécessaires.

Le critère se formule comme suit :

$$\begin{aligned}
 J(\Gamma) = & \int_{\Gamma} k^p(s) \, ds \\
 & + \lambda \int_{\Omega^{in}} k^{in}(\mathbf{x}, \Omega^{in}) \, d\mathbf{x} \\
 & + \epsilon \int_{\Gamma} k^b(s) \, ds
 \end{aligned} \tag{6.19}$$

Les descripteurs k^{in} et k^b sont les mêmes que précédemment et le descripteur k^p est tel que :

$$\begin{cases} k^p &= \varphi(d(\mathbf{x}, \Gamma^{ref})) \\ k^{in} &= \psi(S(\Omega^{in})) \\ k^b &= 1 \end{cases} \tag{6.20}$$

où Γ^{ref} est le contour de référence, $d(\mathbf{x}, \Gamma^{ref})$ la distance signée du point $\mathbf{x} = \Gamma(s)$ au contour de référence, et φ une fonction différentiable.

La dérivée du terme d'a priori $\int_{\Gamma} \varphi(d(\mathbf{x}, \Gamma^{ref})) \, ds$ dans la direction \mathbf{V} est :

$$dJ(\Gamma, \mathbf{V}) = \int_{\Gamma} (-\mathbf{N}^{ref} \cdot \mathbf{N} \, \varphi'(d) - \varphi(d) \, \kappa) \, \mathbf{V} \cdot \mathbf{N} \, ds \tag{6.21}$$

Nous avons choisi comme fonction de pondération de l'a priori : $\varphi = \sqrt{1 + d^2} - 1$. Cette fonction vérifie toutes les conditions émises tout au long du chapitre 5 et rappelées à la section 6.1.

L'équation d'évolution du contour actif est donc :

$$\begin{aligned}
 \frac{\partial \Gamma(p, \tau)}{\partial \tau} = & [\mathbf{N}^{ref} \cdot \mathbf{N} \frac{d}{\sqrt{1 + d^2}} + (\sqrt{1 + d^2} - 1) \, \kappa \\
 & + \lambda \, \psi(S(\Omega^{in}(\tau))) \\
 & + \lambda \, \psi'(S(\Omega^{in}(\tau))) \sum_{i=1}^3 [(I_i(\mathbf{x}) - \mu_i(\Omega^{in}(\tau)))^2 - S(\Omega^{in}(\tau))] \\
 & + \epsilon \kappa] \, \mathbf{N}
 \end{aligned} \tag{6.22}$$

avec $\mathbf{x} = \Gamma(p, \tau)$.

L'algorithme a été mis en œuvre par la méthode des ensembles de niveaux.

6.2.3 Résultats expérimentaux

La figure 6.5 compare les résultats obtenus en utilisant le critère de compétition de régions et ceux obtenus avec l'utilisation du terme d'a priori.

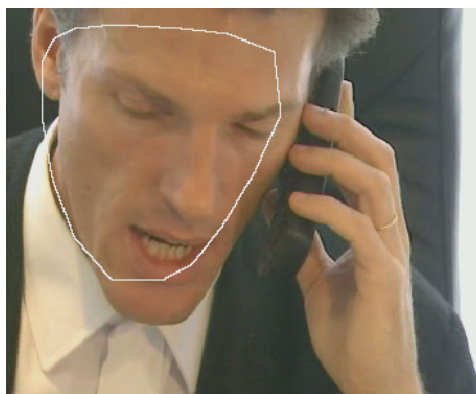
Dans cette image, l'objet d'intérêt est le visage de la personne au téléphone. En utilisant la méthode de compétition de régions, et donc des informations liées à la variance des régions, la main de la personne est considérée comme faisant partie de l'objet car elle a, au sens de l'homogénéité, les mêmes caractéristiques que le visage (figure 6.5.a).

L'introduction du terme d'a priori permet à l'utilisateur d'intervenir pour délimiter grossièrement le centre d'intérêt, en définissant le contour de référence comme sur la figure 6.5.b.

Cette interaction permet d'induire une compétition entre les informations d'homogénéité de la région intérieure au contour actif, et la distance entre ce même contour et le contour de référence défini par l'utilisateur. Ainsi, le contour résultant de ce compromis définira une région la plus homogène possible proche du centre d'intérêt défini par le contour de référence. Avec le terme d'a priori, l'objet final n'inclut plus la main, mais délimite le visage comme souhaité (figure 6.5.c).



(a) Contour initial.



(b) Contour de référence défini par un utilisateur.



(c) Résultat de l'algorithme de compétition de régions : le contour dérive vers la main de la personne au téléphone.



(d) Segmentation avec le terme d'a priori de distance au contour de référence : seul le visage est segmenté, comme souhaité.

FIG. 6.5 – Comparaison entre un algorithme de compétition de régions et l'algorithme de segmentation interactive incluant l'a priori géométrique.

6.3 Segmentation de vidéos

Dans le cadre d'une segmentation d'objets vidéos ou *tracking*, le terme d'a priori permet d'introduire une cohérence temporelle. Cette cohérence est liée au fait que le contour de référence, défini une seule fois par l'utilisateur sur la première image, se déduit ensuite du contour final de l'image précédente, pour les autres images de la séquence.

Nous allons d'abord présenter les étapes spécifiques à la segmentation vidéo, c'est-à-dire principalement l'estimation du contour de référence. L'équation d'évolution du contour actif se déduira ensuite facilement des équations calculées à la section précédente. Enfin, nous montrerons les résultats obtenus sur une séquence de 50 images.

6.3.1 Estimation du contour de référence

Le contour de référence de la première image de la vidéo, noté Γ_0^{ref} est initialisé par un atlas ou interactivement par un opérateur.

Une fois ce premier contour de référence initialisé, tous les autres contours de référence Γ_n^{ref} pour l'image I_n sont estimés à partir du contour Γ_{n-1} , résultat de la segmentation finale de l'image précédente I_{n-1} .

Le contour de référence Γ_n^{ref} se déduit du contour final Γ_{n-1} en tenant compte des différents mouvements entre les deux images I_{n-1} et I_n , dus au mouvement de la caméra et/ou à celui de l'objet.

Pour cela, nous faisons l'hypothèse que le mouvement global du contour entre I_{n-1} et I_n peut être modélisé par un modèle affine à 6 paramètres, $M_{n-1 \rightarrow n}$, ce qui donne :

$$\Gamma_n^{ref} = M_{n-1 \rightarrow n} \Gamma_{n-1} \quad (6.23)$$

Le modèle de mouvement paramétrique affine à 6 paramètres constitue un bon compromis entre la qualité de l'estimation et sa complexité calculatoire.

Une étape préalable à l'évolution du contour actif est donc nécessaire pour définir le contour de référence. Cette définition du contour de référence relie la segmentation de l'image courante au résultat de segmentation de l'image précédente. Elle permet de considérer l'objet d'intérêt comme un objet $2D + t$, et non pas comme une suite de segmentations de n images indépendantes. Cette démarche permet de rectifier les difficultés de segmentation propres à une image particulière en s'appuyant sur les résultats obtenus sur les images précédentes.

Pour estimer le contour de référence Γ_n^{ref} de l'image I_n à partir du contour final Γ_{n-1} on procède comme suit :

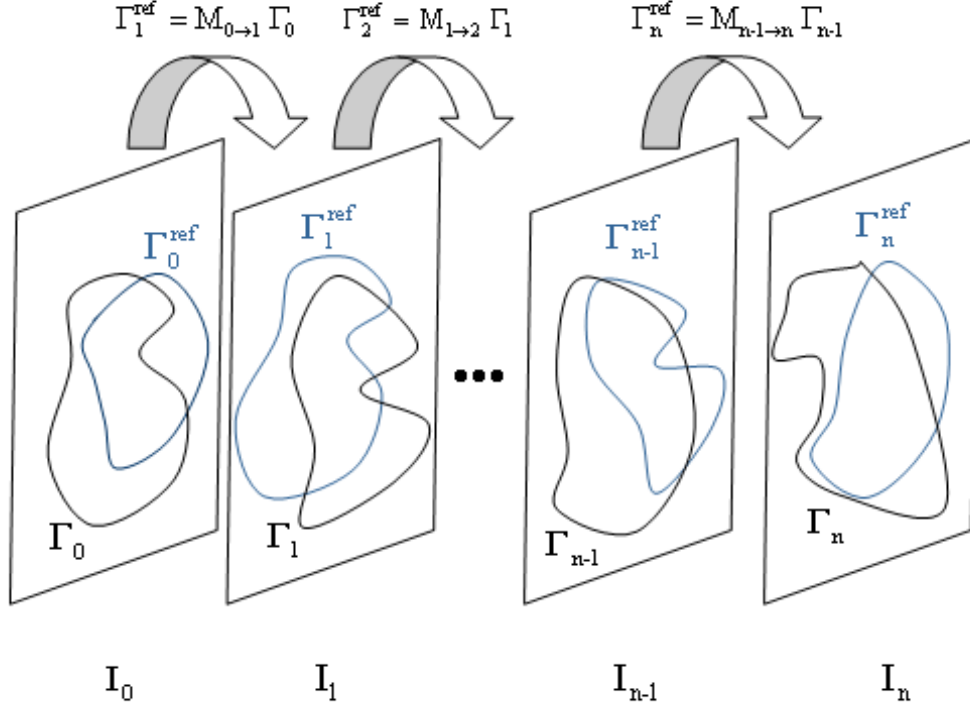


FIG. 6.6 – Estimation du contour de référence pour le tracking

- À chaque pixel \mathbf{x}_{n-1} du contour Γ_{n-1} , nous associons son correspondant \mathbf{x}_n dans l'image I_n à l'aide d'une méthode robuste de mise en correspondance par blocs ou *block-matching*.
- Ensuite, nous cherchons le modèle $M_{n-1 \rightarrow n}$ de mouvement affine représentant au mieux le déplacement global du contour Γ_{n-1} vers I_n , à partir du déplacement de chacun de ses pixels.
Les paramètres du modèle doivent pour cela minimiser la somme des différences pondérées entre les pixels \mathbf{x}_{n-1} , auxquels on applique le modèle de mouvement $M_{n-1 \rightarrow n}$, et leur correspondant \mathbf{x}_n , estimés par la méthode de *block-matching* :

$$\sum_{\mathbf{x}_{n-1} \in \Gamma_{n-1}} \phi(\|\mathbf{x}_n - M_{n-1 \rightarrow n} \mathbf{x}_{n-1}\|) \quad (6.24)$$

La fonction ϕ est la fonction de pénalisation de Geman and McClure [36], elle permet d'améliorer la robustesse de l'estimation en li-

mitant le poids des mises en correspondance aberrantes.

La minimisation de ce critère (6.24) met en oeuvre le théorème semi-quadratique de Charbonnier *et al.* [13], c'est-à-dire que le problème est équivalent à minimiser le critère :

$$\sum_{\mathbf{x}_{n-1} \in \Gamma_{n-1}} (b_{min} \|\mathbf{x}_n - M_{n-1 \rightarrow n} \mathbf{x}_{n-1}\|^2 + \psi(b_{min})) \quad (6.25)$$

avec

$$b_{min} = \frac{\phi'(\|\mathbf{x}_n - M_{n-1 \rightarrow n} \mathbf{x}_{n-1}\|)}{2\|\mathbf{x}_n - M_{n-1 \rightarrow n} \mathbf{x}_{n-1}\|} \quad (6.26)$$

La minimisation du critère (6.24) se fait en calculant alternativement le modèle de mouvement $M_{n-1 \rightarrow n}$ à partir de l'équation (6.25) pour b_{min} fixé, et en calculant la carte des b_{min} à partir de l'équation (6.26) pour $M_{n-1 \rightarrow n}$ fixé.

L'estimation robuste des paramètres d'un modèle de mouvement suivant cette méthode est présenté plus en détails en Annexe A.1.2 page 138.

- Enfin le contour Γ_{n-1} est projeté dans l'image I_n suivant le modèle de mouvement $M_{n-1 \rightarrow n}$ et définit le contour de référence de l'image I_n soit : $\Gamma_n^{ref} = M_{n-1 \rightarrow n} \Gamma_{n-1}$.

6.3.2 Critère de segmentation de vidéos et équation d'évolution du contour

Le critère de segmentation de vidéos J de l'image I_n de la vidéo comprend, comme pour le critère de segmentation d'image, un terme d'a priori au contour de référence, un terme d'attache aux propriétés de l'objet et un terme de régularisation du contour.

$$\begin{aligned} J(\Gamma_n) = & \int_{\Gamma_n} k_n^p(s, \Gamma_n^{ref}) ds \\ & + \lambda \int_{\Omega_n^{in}} k_n^{in}(\mathbf{x}, \Omega_n^{in}) d\mathbf{x} \\ & + \epsilon \int_{\Gamma_n} k^b(s) ds \end{aligned} \quad (6.27)$$

Le descripteur k_n^p représente la distance pondérée au contour de référence Γ_n^{ref} , estimé à partir du contour final de l'image précédente, comme décrit à la section (6.3.1).

Le descripteur de l'objet k_n^{in} est un descripteur statistique de régions homogènes.

Le terme de régularisation du contour minimise sa longueur. Enfin λ et ϵ sont des constantes de pondération utilisées pour faire varier l'influence de chacun des termes sur le critère.

$$\begin{cases} k_n^p &= \varphi(d(\mathbf{x}, \Gamma_n^{ref})) \\ k_n^{in} &= \psi(S(\Omega_n^{in})) \\ k^b &= 1 \end{cases} \quad (6.28)$$

où Γ_n^{ref} est le contour de référence, $d(\mathbf{x}, \Gamma_n^{ref})$ la distance signée d'un point $\mathbf{x} = \Gamma_n(s)$ du contour actif Γ_n au contour de référence, et φ une fonction différentiable.

La description du terme de région nécessite l'introduction de quelques notations :

- $I^i(\mathbf{x})$ représente la luminance ($i = 1$) ou l'une des deux chrominances ($i = 2, 3$) de l'image I au pixel \mathbf{x} ,
 - $\mu_i(\Omega)$ est la valeur moyenne de I pour la composante couleur i sur la région $\Omega = \Omega^{in}$ ou Ω^{out} ,
 - $\sigma_i^2(\Omega)$ est la variance de I pour la composante couleur i sur la région $\Omega = \Omega^{in}$ ou Ω^{out} ,
 - $S(\Omega) = \sigma_1^2(\Omega) + \sigma_2^2(\Omega) + \sigma_3^2(\Omega)$ est la somme des variances de chaque composante couleur de I sur la région Ω .
 - et ψ est une fonction de pondération différentiable.
- Nous avons choisi $\psi = \log(1 + r^2)$ ($\psi' = \frac{2r}{1+r^2}$).

L'équation d'évolution se déduit des résultats présentés à la section 6.2.2 :

$$\begin{aligned} \frac{\partial \Gamma_n(p, \tau)}{\partial \tau} &= [\mathbf{N}^{ref} \cdot \mathbf{N} \frac{d}{\sqrt{1+d^2}} + (\sqrt{1+d^2} - 1)\kappa \\ &+ \lambda \psi(S(\Omega_n^{in}(\tau))) \\ &+ \lambda \psi'(S(\Omega_n^{in}(\tau))) \sum_{i=1}^3 [(I^i(\mathbf{x}) - \mu_i(\Omega_n^{in}(\tau)))^2 - S(\Omega_n^{in}(\tau))] \\ &+ \epsilon \kappa] \mathbf{N} \end{aligned} \quad (6.29)$$

où $\mathbf{x} = \Gamma_n(p, \tau)$.

6.3.3 Résultats expérimentaux

Les résultats présentés dans la figure 6.7 illustrent le suivi d'un visage le long d'une séquence de 50 images.

Le premier contour de référence est défini interactivement comme sur la figure 6.7(a).

Le pas d'évolution du contour actif est : $\delta\tau = 5.10^{-4}$, et les constantes de pondération entre l'a priori, le terme d'attache à la région et la régularisation du contour valent : $\lambda = 0.6$ et $\epsilon = 115$.



(a) Image 0



(b) Image 9



(c) Image 19



(d) Image 29



(e) Image 39



(f) Image 49

FIG. 6.7 – Segmentation de visage utilisant l'a priori géométrique sur 50 images de la vidéo Erik.

Troisième partie

Segmentation et estimation conjointes du mouvement

Chapitre 7

État de l'art

La segmentation d'objets en mouvement joue un rôle de plus en plus important dans les applications multimédia, notamment depuis l'apparition des normes de compression vidéo MPEG-4 [40], MPEG-7 [41] et H264, dites de seconde génération.

De plus, la segmentation d'objets en mouvement s'étend souvent au suivi de ces objets ou *tracking* le long d'une séquence vidéo. Les domaines d'applications dans ce cas touchent aussi bien le domaine médical [56] que la télé-surveillance [90], la rotoscopie [32], la météorologie [68] ou l'indexation [59].

En compression vidéo, les normes MPEG-1 [42] et MPEG-2 [43], H261 [45] et H263 [46] exploitent les redondances temporelles pour réduire le nombre de données à transmettre. En effet, si l'on néglige les éventuels changements d'intensité lumineuse et le bruit d'acquisition, la majeure partie des pixels ne varie pas d'une image à l'autre, pour une vidéo à caméra fixe. Dans le cas d'une séquence à caméra mobile, il est possible d'estimer et de compenser le mouvement de la caméra pour se rapprocher d'un cas de caméra fixe.

La réduction de la quantité d'information à transmettre s'obtient donc en transmettant l'image différence entre l'image courante et l'image suivante compensée en mouvement, connue sous le nom de DFD (*displaced frame difference*), ainsi que les vecteurs de mouvement correspondants.

Une des étapes-clés de ces normes réside dans l'estimation du mouvement, estimation nécessaire pour la compensation du mouvement de la caméra notamment. La méthode d'estimation du mouvement la plus répandue dans ce cadre est le *block-matching* [53, 92].

Cependant, l'utilisation du *block-matching* induit un découpage arbitraire de l'image en blocs, arbitraire dans le sens où le découpage ne prend pas en compte le contenu sémantique de la vidéo. Un bloc peut contenir à la fois des pixels du fond de l'image et des pixels d'un objet, et donc inclure

des régions ayant des mouvements différents. Sur ces blocs l'estimation est mauvaise, ce qui retentit sur la qualité de la compensation et celle de la DFD.

Les codeurs de seconde génération MPEG-4 [40] MPEG-7 [41] et H264 ont opté pour une approche sémantique de la vidéo. Le découpage de la vidéo ne se fait plus en blocs mais en objets vidéo, ou VOP (*Video Object Plan*). Cette nouvelle représentation de la séquence autorise de nouvelles fonctionnalités, comme interagir avec les objets de l'image ou coder différemment les objets et le fond de la séquence, ou indexer les objets d'une vidéo et rechercher leur présence dans d'autres.

Dans le cadre de ces nouvelles applications multimédia, la segmentation des objets est donc une étape-clé. Pour segmenter des objets, il convient tout d'abord de définir les notions d'objets, et donc par complémentarité, de fond de l'image.

En l'absence de mouvement de caméra, les seules différences entre deux images d'une séquence sont dues aux objets en mouvement. Partant de ce constat simple, une segmentation courante de l'image consiste à définir le fond de l'image comme la partie statique et les objets comme la partie mobile de l'image, ce qui revient à détecter le mouvement. En pratique, il faudra prendre en compte dans notre définition un éventuel mouvement de la caméra entre les images considérées.

La détection du mouvement ne nécessite pas l'estimation du mouvement des objets. Cependant la détection des objets en mouvement ne permet pas, sans autre hypothèse, ni de différencier les objets en mouvement s'il y en a plusieurs, ni de connaître leur mouvement. Or seule la connaissance du mouvement de l'objet permet son suivi (ou *tracking*) le long de la séquence vidéo, suivi nécessaire à la définition du VOP.

Nous allons introduire quelques travaux présentant des méthodes de détection de mouvement, puis de segmentation et suivi d'objets en mouvement avant de situer notre approche dans ce contexte.

7.1 Détection de mouvement

La détection de mouvement permet de séparer l'image en deux régions :

- une région **statique** : le **fond** de l'image,
- une région en **mouvement** : le ou les **objets**.

En pratique, il faut distinguer les séquences à caméra fixe et les séquences à caméra mobile, pour lesquelles le mouvement de la caméra doit être com-

pensé. De plus, certaines méthodes de détection n'utilisent que deux images successives, alors que d'autres méthodes étendent leur support temporel à plusieurs images, en estimant une image représentant le fond statique de la scène. Nous allons évoquer quelques méthodes relatives à ces différents cas.

Une des méthodes les plus simples de détection de mouvement **entre deux images à caméra fixe** se fonde sur la différence entre ces deux images, la FD (*frame difference*). La différence entre les images s'opère pixel à pixel et l'image résultat est seuillée. Si la différence dépasse le seuil fixé, le pixel est considéré comme appartenant à un objet en mouvement, sinon il est considéré comme appartenant au fond de l'image.

Cette méthode a pour avantage de ne pas nécessiter l'estimation du mouvement des objets. Elle est cependant sensible aux variations d'intensité lumineuse et au bruit. De plus, les zones découvertes par l'objet d'une image à l'autre sont étiquetées comme faisant partie de l'objet.

La FD peut également servir de base à une approche statistique du problème de détection [70]. La densité de probabilité de la FD est considérée comme un mélange d'une composante statique et d'une composante mobile.

Dans le cas de séquences à **caméra mobile**, on ne peut plus définir l'objet comme l'ensemble des pixels en mouvement car le mouvement de la caméra induit un mouvement dominant dans la séquence sur tous les pixels de l'image. Cependant les pixels de l'objet ont un mouvement local propre qui s'ajoute à ce mouvement global. L'objet va donc être défini comme l'ensemble des pixels ne suivant pas le mouvement dominant de l'image.

Une méthode simple consiste à compenser le mouvement de la caméra, c'est-à-dire recalculer l'une des images sur l'autre, pour se replacer dans un cadre de caméra fixe. La différence entre les deux images, dont l'une est compensée en mouvement, est appelée la DFD (*Displaced frame difference*). Il est donc nécessaire d'estimer le mouvement de la caméra entre les deux images pour pouvoir le compenser et discriminer les pixels ne suivant pas ce mouvement. Le mouvement de la caméra est souvent représenté par un modèle de mouvement, dont l'estimation tend à être la plus robuste possible [13, 66].

Wu et Kittler [89] représentent le mouvement de la caméra par un modèle affine. L'estimation des paramètres du modèle est formulée comme la minimisation de la valeur absolue de la TFD (*Transformed Frame Difference*),

image différence entre l'image courante et l'image suivante transformée par les paramètres de mouvement. Enfin les objets sont détectés par seuillage de la valeur absolue de la TFD.

Dans [77] Salembier *et al.* proposent de représenter l'image sous forme d'arbre. Pour une segmentation d'objets en mouvement, la DFD est calculée sur l'image entière, considérée comme la racine de l'arbre, et définit deux régions (deux enfants) pour lesquels le processus de détection est itéré. En choisissant un niveau de description de l'arbre, on détermine le nombre de régions qui segmentent l'image.

Precioso *et al.* [72] utilisent un algorithme de contours actifs mis en œuvre par des *smoothing B-Splines*. Le contour actif évolue de sorte à minimiser la différence sur le fond de l'image entre l'image courante et la précédente compensée en mouvement. Afin d'éviter la sur-segmentation de l'objet le critère de segmentation inclut un terme visant à minimiser l'aire de l'objet et un terme de régularisation portant sur la longueur du contour.

Fablet *et al.* [29] proposent une méthode intégrant des informations spatiales sous la forme d'une segmentation basée sur la couleur. Un modèle de mouvement affine est estimé afin de modéliser au mieux le mouvement dominant de l'image. Enfin chaque région déterminée par la segmentation spatiale est étiquetée comme appartenant au fond ou à l'objet respectivement si elle suit ou non le mouvement dominant.

Dans [85] la détection de mouvement repose sur le principe de groupement perceptuel qui stipule que les événements perceptibles sont ceux qui diffèrent d'un modèle aléatoire général. Il s'agit donc de modéliser la partie statique de la scène vide d'objets en mouvement afin de détecter *a contrario* ces objets.

Toujours dans l'optique de séparer le fond de l'image des objets en mouvement, certaines méthodes se sont penchées sur l'estimation d'une image représentant la partie statique de la séquence que nous appellerons «fond de la séquence». Le fond de la séquence étant connu, la détection des objets en mouvement est directe. Ces méthodes utilisent à cette fin **plusieurs images** successives de la séquence. Là encore il est nécessaire de distinguer les séquences à caméra fixe des séquences à caméra mobile.

Pour des séquences à **caméra fixe**, Jehan-Besson *et al.* [50] proposent d'estimer le fond d'une séquence de n images comme une moyenne de ces n images. Afin d'éliminer les valeurs correspondant aux objets en mouvement,

les auteurs introduisent dans leur critère un M-estimateur et minimisent la différence pondérée entre le fond et chacune des images par un algorithme de minimisations alternées.

Dans [54] les auteurs couplent l'estimation du fond de la séquence à sa restauration.

Pour des séquences à **caméra mobile** le fond de la séquence est appelé mosaïque. De nombreuses méthodes de construction de mosaïques sont présentées dans [44].

L'annexe A développe une méthode variationnelle de contours actifs pour la détection de mouvement présentée dans [33]. Le critère de segmentation est fondé sur la connaissance du fond de la séquence sous forme de mosaïque. La construction de la mosaïque nécessite une pré-segmentation grossière des objets. En utilisant plusieurs images de la séquence, la détection gagne en robustesse par rapport au bruit et aux changements d'intensité lumineuse.

7.2 Segmentation et suivi d'objets en mouvement

Si la détection d'objets en mouvement permet de segmenter l'image entre régions en mouvement et région statique, elle ne donne aucune information quant au mouvement de ces régions, or cette information est nécessaire pour des applications de type *tracking*.

Parmi les méthodes de segmentation et de *tracking* d'objets en mouvement nous présenterons deux grandes familles : celles dites séquentielles, qui segmentent l'image à partir de données calculées au préalable, et celles, couplées ou conjointes, qui lient la segmentation et l'estimation du mouvement dans un procédé itératif.

Enfin nous introduirons nos travaux et les situerons parmi les méthodes d'estimation et de segmentation conjointes du mouvement.

7.2.1 Segmentation et estimation séquentielle du mouvement

Certaines méthodes estiment le flot optique sur l'ensemble de l'image considérée, et partitionnent ensuite l'image à partir de ce champ dense.

Par exemple, les méthodes [37, 68, 73, 76] mettent en œuvre un algorithme de segmentation par contours actifs utilisant l'estimation préalable d'un champ de mouvement dense.

Dans [73] les auteurs définissent un critère qui mêle entre autres un terme fondé sur la norme des vecteurs de mouvement et un terme d'attraction vers les discontinuités d'intensité lumineuse.

Dans le cadre d'applications météorologiques, Papin *et al.* [68] proposent d'associer des données photométriques à une estimation robuste d'un champ de vecteurs de mouvement dense.

Herbulot *et al.* [37] définissent un critère de segmentation de données vectorielles exploitant l'entropie conjointe des vecteurs de mouvement.

Dans [76] les auteurs segmentent un champ de vecteurs pouvant être décrit par un nombre fini de paramètres représentant la trajectoire des pixels de l'objet.

Ces méthodes ont en commun l'estimation d'un champ dense de vecteurs de mouvement préalable à l'algorithme de segmentation.

Parmi les approches de segmentation nécessitant un pré-traitement certaines n'utilisent pas une estimation du flot optique à proprement parler

L'algorithme de *tracking* présenté dans [61] détecte la surface spatio-temporelle générée par les frontières dues au mouvement. Le terme définissant la frontière due au mouvement s'appuie sur l'équation du flot optique et l'hypothèse que le mouvement est localement constant partout sauf à ses frontières. Le critère de segmentation inclut également un terme statistique définissant les densités de probabilités des objets en mouvement et du fond de l'image en fonction de la composante normale du flot optique, sans toutefois calculer explicitement celui-ci.

7.2.2 Segmentation et estimation conjointe du mouvement

La performance des méthodes de segmentation basées sur l'estimation préalable du flot optique dépend fortement de la qualité d'estimation du flot optique. Puisque l'on cherche une région ayant un mouvement cohérent, il est intéressant d'utiliser les informations propres à la région ainsi déterminée pour calculer son mouvement. L'estimation du mouvement sera d'autant plus juste que la région segmentée sera proche de l'objet d'intérêt et réciproquement la segmentation sera d'autant plus précise que l'estimation du mouvement sera proche du mouvement apparent réel. Le problème se pose ainsi en un procédé itératif.

Les méthodes d'estimation et de segmentation conjointes du mouvement définissent un critère dépendant à la fois de la région et du mouvement à estimer. La minimisation de ce critère conjoint a pu être traité comme une minimisation hiérarchique non linéaire [62], comme un problème aux

valeurs propres [21] ou par l'évolution d'un contour actif [7, 25, 34, 56, 70, 79].

Les contours actifs sont particulièrement bien adaptés pour la segmentation et le suivi d'objets en mouvement [7, 25, 34, 56, 70, 79] parce que les objets bougent et se déforment peu entre deux images. De ce fait il est fréquent d'initialiser la segmentation d'une image par le contour final de l'image précédente, et l'algorithme converge d'autant plus vite que le contour initial est proche de l'objet d'intérêt.

Dans [89] les auteurs formulent l'hypothèse d'un modèle de mouvement global pour toute la région d'intérêt. Les auteurs utilisent la valeur absolue de l'image différence.

Dans [20, 89] les auteurs proposent un critère d'estimation et de segmentation en mouvement, et minimisent successivement l'énergie par rapport à une des deux inconnues : la région d'intérêt ou les paramètres du mouvement, en fixant l'autre inconnue. Dans ce cas l'estimation du mouvement et la segmentation sont couplées.

Segmentation et estimation conjointes du mouvement entre deux images successives

Dans ce chapitre, nous définissons un critère de segmentation d'un mouvement, conjointement à l'estimation du mouvement de la région segmentée, à partir de deux images successives. Nous dérivons le critère avec la méthode des gradients de forme, et proposons deux équations d'évolution : une dirigée suivant la normale au contour, pour la segmentation d'un objet en mouvement, et une dirigée suivant le mouvement de l'objet pour le suivi d'objets en mouvement. Nous étudions la méthode d'estimation du mouvement ainsi définie pour un modèle de mouvement polynômial, en nous attachant plus particulièrement aux modèles de translation et d'homothétie.

8.1 Définition du critère conjoint entre deux images

Nous proposons une méthode de segmentation et d'estimation d'un mouvement utilisant les contours actifs. Dans cette section nous définirons notre critère de segmentation et les hypothèses sous-jacentes, tout en situant nos travaux par rapport à l'existant.

Notons I une séquence d'images I_n en niveaux de gris. La séquence I peut être vue comme une fonction de $S \equiv D \times [0, N]$ dans \mathcal{R}^+ , où D est un ouvert borné et N un entier positif. Nous noterons $I_n(\mathbf{x})$ le réel positif $I(\mathbf{x}, n)$ pour tous \mathbf{x} et n .

Une image est une projection dans le plan $2D$ d'une scène réelle. L'estimation du mouvement à partir d'une séquence ne peut déterminer le mouvement

"réel" des objets 3D qui composent la scène, mais seulement leur mouvement apparent. On admet communément l'hypothèse que l'intensité lumineuse d'un pixel ne varie pas le long de sa trajectoire [39, 2, 88]. Le mouvement apparent est donc calculé à partir des constances de l'intensité lumineuse d'une image à l'autre.

Soit \mathbf{v} le flot optique entre les images I_n et I_{n-1} , c'est-à-dire le champ de vecteurs représentant le mouvement apparent entre ces deux images. L'hypothèse de constance de l'intensité s'écrit :

$$I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + \mathbf{v}(\mathbf{x})) = 0 \quad (8.1)$$

Ainsi posé le problème d'estimation du flot optique est sous-déterminé et $\mathbf{v}(\mathbf{x})$ n'est pas unique, car de nombreux pixels ont la même intensité lumineuse.

Afin de contraindre l'équation (8.1), nous posons l'hypothèse d'un mouvement constant par région, en étendant le domaine d'application de l'équation à un voisinage Ω_n du point considéré :

$$\int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + \mathbf{v}(\mathbf{x})))^2 = 0 \quad (8.2)$$

Cette hypothèse se retrouve, entre autres, dans les travaux de [89] et dans la méthode de *block-matching* [42, 43] pour le cas particulier où la région de support est un carré.

Nous restreignons également le mouvement en le représentant par un modèle de mouvement paramétrique [20, 66]. Le mouvement de la région Ω_n est modélisé par un vecteur $\mathbf{p}(\Omega_n)$ de m paramètres. Le vecteur $\mathbf{p}(\Omega_n)$ est constant sur la région Ω_n , ceci pour respecter l'hypothèse de constance du mouvement par région :

$$\mathbf{v}(\mathbf{x}) = M(\mathbf{x}, n) \mathbf{p}(\Omega_n), \quad \mathbf{x} \in \Omega_n \quad (8.3)$$

M est une matrice $2 \times m$ dépendant du modèle de mouvement choisi. Par exemple, si le modèle de mouvement choisi est la translation uniforme de vecteur \mathbf{a} , la matrice $M(\mathbf{x}, n)$ est la matrice identité et le vecteur des paramètres du mouvement $\mathbf{p}(\Omega_n)$ est le vecteur de translation \mathbf{a} .

A partir des hypothèses décrites dans les équations (8.2) et (8.3), nous définissons l'énergie suivante (voir figure 8.1) :

$$E(\Omega_n) = \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M(\mathbf{x}, n) \mathbf{p}(\Omega_n)))^2 d\mathbf{x} \quad (8.4)$$

où $\mathbf{p}(\Omega_n)$ représente le vecteur des paramètres de mouvement de la région Ω_n .

Nous avons choisi les paramètres de mouvement $\mathbf{p}(\Omega_n)$ tels que :

$$\mathbf{p}(\Omega_n) = \min_{\mathbf{p}} \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M(\mathbf{x}, n) \mathbf{p}))^2 d\mathbf{x} \quad (8.5)$$

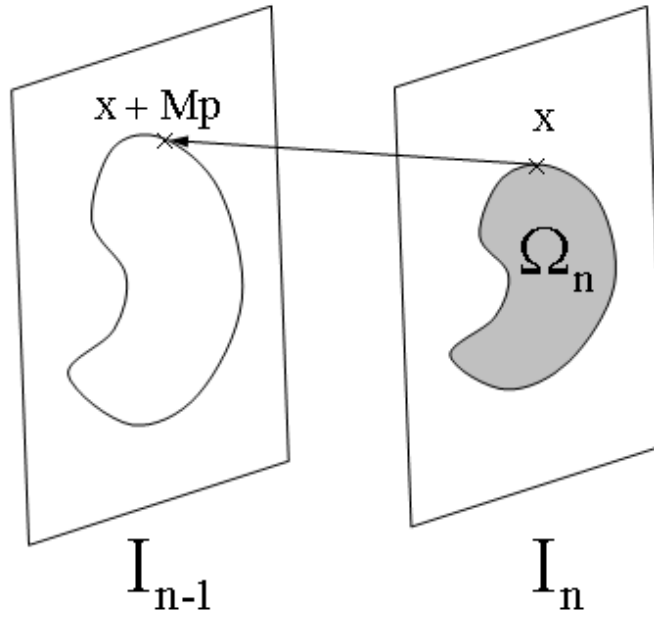


FIG. 8.1 – Modèle de mouvement entre deux images

Minimiser l'énergie (8.4) par rapport à Ω_n revient à délimiter une région sur laquelle le mouvement est homogène et caractérisé par un modèle de mouvement, de fait constant sur la région, et défini par l'équation (8.5).

Dans [89] les auteurs étudient la valeur absolue de la différence entre deux images successives. Cependant la valeur absolue n'est pas différentiable en zéro et pose problème lors de la différentiation du critère, différentiation nécessaire au calcul de l'équation d'évolution du contour actif.

Contrairement à l'approche couplée, développée par [20, 89], notre approche de l'estimation du mouvement et de la segmentation est jointe, car nous prenons explicitement en compte, lors de la minimisation de l'énergie, le fait que les paramètres $\mathbf{p}(\Omega_n)$ du modèle de mouvement dépendent de la

région Ω_n (voir section 8.1).

La segmentation de l'objet en mouvement s'effectue par contours actifs. Le minimum de l'énergie (8.4) est atteint lorsque la région Ω_n recouvre l'objet en mouvement. Ainsi, partant d'un contour initial, le contour actif doit évoluer de sorte à faire décroître le critère (8.4). Nous déduisons donc l'équation d'évolution du contour actif à partir de la dérivée de ce critère.

L'estimation du mouvement de l'objet passe par l'estimation des paramètres du mouvement, définis dans l'équation (8.5) comme argument minimum de la fonctionnelle présentée.

Nous verrons par la suite que le choix de la fonctionnelle à minimiser pour l'estimation des paramètres du mouvement influe sur l'équation d'évolution du contour actif, et donc sur le résultat de la segmentation (Démonstration 8.2 page 93).

8.2 Segmentation et suivi d'objets en mouvement

La minimisation du critère (8.4) requiert le calcul de sa dérivée par rapport à Ω_n , calcul effectué en utilisant les outils de dérivation de domaines.

On utilise un schéma dynamique dans lequel la région Ω_n dépend d'une variable d'évolution τ . La dérivée du critère dans la direction \mathbf{V} est calculée suivant la technique des gradients de forme [26, 51].

Théorème 8.1. *La dérivée eulérienne dans la direction \mathbf{V} du critère*

$$E(\Omega_n) = \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M(\mathbf{x}, n) \mathbf{p}(\Omega_n)))^2 d\mathbf{x} \quad (8.6)$$

où les paramètres du modèle de mouvement $\mathbf{p}(\Omega_n)$ vérifient :

$$\mathbf{p}(\Omega_n) = \min_{\mathbf{p}} \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M(\mathbf{x}, n) \mathbf{p}))^2 d\mathbf{x} \quad (8.7)$$

à pour expression :

$$dE(\Omega_n, \mathbf{V}) = - \int_{\Gamma_n} (I_n(s) - I_{n-1}(s + M \mathbf{p}(\Omega_n)))^2 \mathbf{V} \cdot \mathbf{N} ds \quad (8.8)$$

Dans cette expression, Γ_n représente le contour orienté de la région Ω_n et s son abscisse curviligne. Nous noterons $s + M \mathbf{p}(\Omega_n)$ l'abscisse du point \mathbf{x}_{n-1}

de Γ_{n-1} tel que $\mathbf{x}_{n-1} = \mathbf{x} + M(\mathbf{x}, n) \mathbf{p}(\Omega_n)$, où $\mathbf{x} = \Gamma_n(s)$. La déformation locale du contour (sa vitesse) est notée \mathbf{V} et le vecteur \mathbf{N} représente la normale unitaire intérieure au contour Γ_n .

Dans le cadre d'une segmentation, une équation d'évolution possible est (voir section 8.2.1) :

$$\frac{\partial \Gamma_n}{\partial \tau}(p, \tau) = \alpha(I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n(\tau))))^2 \mathbf{N} \quad (8.9)$$

avec $\mathbf{x} = \Gamma_n(p, \tau)$.

Nous présentons à la section 8.2.2 une autre équation d'évolution spécifique aux applications de suivi d'objets en mouvement.

Démonstration :

Le théorème 3.1 page 25 nous permet d'exprimer la dérivée eulérienne dans la direction \mathbf{V} du critère en fonction de sa dérivée de domaine :

$$\begin{aligned} dE(\Omega_n, \mathbf{V}) &= \int_{\Omega_n} d_{\Omega}((I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n))))^2(\mathbf{x}, \Omega_n, \mathbf{V}) \, d\mathbf{x} \\ &\quad - \int_{\Gamma_n} (I_n(s) - I_{n-1}(s + M \mathbf{p}(\Omega_n)))^2 \mathbf{V} \cdot \mathbf{N} \, ds \end{aligned} \quad (8.10)$$

Le terme région de cette équation est noté A , soit :

$$A = \int_{\Omega_n} d_{\Omega}((I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n))))^2(\mathbf{x}, \Omega_n, \mathbf{V}) \, d\mathbf{x} \quad (8.11)$$

Développons la dérivée partielle contenue dans le terme région A comme une dérivée de fonctions composées.

$$A = \int_{\Omega_n} d_{\Omega} \mathbf{p}(\mathbf{x}, \Omega_n, \mathbf{V}) \frac{\partial}{\partial \mathbf{p}} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n)))^2 \, d\mathbf{x} \quad (8.12)$$

Avec, par hypothèse, un modèle de mouvement $\mathbf{p}(\Omega_n)$ unique pour toute la région Ω_n , on en déduit donc les égalités suivantes :

$$A = d_{\Omega} \mathbf{p}(\mathbf{x}, \Omega_n, \mathbf{V}) \int_{\Omega_n} \frac{\partial}{\partial \mathbf{p}} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n)))^2 \, d\mathbf{x} \quad (8.13)$$

$$= d_{\Omega} \mathbf{p}(\mathbf{x}, \Omega_n, \mathbf{V}) \frac{\partial}{\partial \mathbf{p}} \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n)))^2 \, d\mathbf{x} \quad (8.14)$$

Ce qui donne :

$$A = d_{\Omega} \mathbf{p}(\mathbf{x}, \Omega_n, \mathbf{V}) \frac{\partial}{\partial \mathbf{p}} G(\mathbf{p}(\Omega_n), \Omega_n) \quad (8.15)$$

avec

$$G(\mathbf{p}(\Omega_n), \Omega_n) = \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n)))^2 d\mathbf{x} \quad (8.16)$$

Sans autre hypothèse sur les paramètres de mouvement, la dérivée eulérienne du critère de segmentation est :

$$\begin{aligned} dE(\Omega_n, \mathbf{V}) &= d_{\Omega} \mathbf{p}(\mathbf{x}, \Omega_n, \mathbf{V}) \frac{\partial}{\partial \mathbf{p}} G(\mathbf{p}(\Omega_n), \Omega_n) \\ &- \int_{\Gamma_n} (I_n(s) - I_{n-1}(s + M \mathbf{p}(\Omega_n)))^2 \mathbf{V} \cdot \mathbf{N} ds \end{aligned} \quad (8.17)$$

Cette dérivée ne nous permet pas d'en déduire directement une équation d'évolution pour le contour actif, car elle ne se présente pas sous la forme d'une seule intégrale de contour.

Lemme 8.1. *Si les paramètres $\mathbf{p}(\Omega_n)$ sont définis par*

$$\mathbf{p}(\Omega_n) = \min_{\mathbf{p}} \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}))^2 d\mathbf{x} \quad (8.18)$$

alors le terme région A est nul :

$$A = \int_{\Omega_n} d_{\Omega} ((I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n)))^2)(\mathbf{x}, \Omega_n, \mathbf{V}) d\mathbf{x} = 0 \quad (8.19)$$

Démonstration :

Les paramètres de mouvement ont été choisis tels que :

$$\begin{aligned} \mathbf{p}(\Omega_n) &= \min_{\mathbf{p}} \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}))^2 d\mathbf{x} \\ &= \min_{\mathbf{p}} G(\mathbf{p}, \Omega_n) \end{aligned} \quad (8.20)$$

Sans contrainte sur \mathbf{p} ($\mathbf{p} \in R^m$), comme les paramètres de mouvement $\mathbf{p}(\Omega_n)$ vérifient le minimum de G , la dérivée par rapport à \mathbf{p} de G évaluée au point $\mathbf{p}(\Omega_n)$ est nulle :

$$\frac{\partial}{\partial \mathbf{p}} G(\mathbf{p}(\Omega_n), \Omega_n) = 0 \quad (8.21)$$

Dans le cas où les paramètres de mouvement sont contraints (par exemple, $\mathbf{p} \in R^{m-1} \times R^+$), nous faisons l'hypothèse que ces contraintes peuvent s'exprimer par $\phi(\mathbf{p}) = 0$.

Il existe alors un multiplicateur de Lagrange λ tel que

$$\frac{\partial}{\partial \mathbf{p}} G(\mathbf{p}(\Omega_n), \Omega_n) = \lambda \phi'(\mathbf{p}(\Omega_n)) \quad (8.22)$$

Le terme région A s'écrit :

$$A = \lambda d_{\Omega} \mathbf{p}(\mathbf{x}, \Omega_n, \mathbf{V}) \phi'(\mathbf{p}(\Omega_n)) \quad (8.23)$$

$$= \lambda d \mathbf{p}(\Omega_n, \mathbf{V}) \phi'(\mathbf{p}(\Omega_n)) \quad (8.24)$$

$$= \lambda d(\phi(\mathbf{p}(\Omega_n)))(\Omega_n, \mathbf{V}) \quad (8.25)$$

Comme l'expression de la contrainte sur \mathbf{p} s'exprime par $\phi(\mathbf{p}) = 0 \quad \forall \mathbf{p}$, ϕ ne varie pas en fonction de Ω_n , et l'expression $d(\phi(\mathbf{p}(\Omega_n)))(\Omega_n, \mathbf{V})$ est nulle. \square

Dans tous les cas, l'intégrale sur la région de la dérivée de domaine s'annule :

$$\begin{aligned} A &= \int_{\Omega_n} d_{\Omega}((I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n)))^2)(\mathbf{x}, \Omega_n, \mathbf{V}) d\mathbf{x} \\ &= d_{\Omega} \mathbf{p}(\mathbf{x}, \Omega_n, \mathbf{V}) \frac{\partial}{\partial \mathbf{p}} G(\mathbf{p}(\Omega_n)) \\ &= 0 \end{aligned} \quad (8.26)$$

et la dérivée du critère ne contient qu'un terme contour :

$$dE(\Omega_n, \mathbf{V}) = - \int_{\Gamma_n} (I_n(s) - I_{n-1}(s + M \mathbf{p}(\Omega_n)))^2 \mathbf{V} \cdot \mathbf{N} ds \quad (8.27)$$

\square

8.2.1 Équation d'évolution pour la segmentation d'objets en mouvement

Un contour initial est itérativement déformé suivant une vitesse \mathbf{V} telle que la dérivée (8.8) du critère soit négative ou nulle. Ainsi, nous nous assurons que le contour évolue vers un minimum local ou le minimum global du critère, s'il existe. Ce minimum est atteint lorsque la dérivée est nulle, et le contour correspondant est alors la segmentation finale de l'image.

L'équation d'évolution du contour, déduite de la dérivée du critère, est :

$$\frac{\partial \Gamma_n}{\partial \tau}(p, \tau) = \alpha(I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n(\tau))))^2 \mathbf{N} \quad (8.28)$$

La constante positive α est un pas d'évolution et $\mathbf{x} = \Gamma_n(p, \tau)$.

Notons que \mathbf{V} est dirigée suivant la normale de Γ_n .

Caselles *et al.* [9] ont démontré que la composante tangentielle de la vitesse n'influe pas sur la propagation du contour actif, mais uniquement sur sa paramétrisation.

Remarquons, de plus, que la déformation du contour nécessite le calcul des paramètres du mouvement à chaque itération (cf. section 8.3).

8.2.2 Équation d'évolution pour le suivi d'objets en mouvement

En général, les algorithmes de suivi d'objets en mouvement [34, 56, 70, 79] choisissent comme contour initial $\Gamma_n(\tau = 0)$ pour l'image courante le contour final de la segmentation de l'image précédente. L'équation d'évolution du contour actif est alors la même que pour une segmentation classique (par exemple Eq. (8.28)) sans aucune autre particularité caractérisant l'objet en mouvement.

Cependant, il peut être intéressant d'utiliser le fait que le contour initial détoure l'objet d'intérêt dans l'image I_{n-1} , et que le contour final doit être l'objet dans l'image suivante I_n .

C'est pourquoi nous avons défini une nouvelle équation d'évolution qui prendrait en compte le mouvement estimé de l'objet, $M \mathbf{p}(\Omega_n)$ entre ces deux images :

$$\frac{\partial \Gamma_n}{\partial \tau}(p, \tau) = \alpha(I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega_n(\tau))))^2 M \mathbf{p}(\Omega_n(\tau)) \quad (8.29)$$

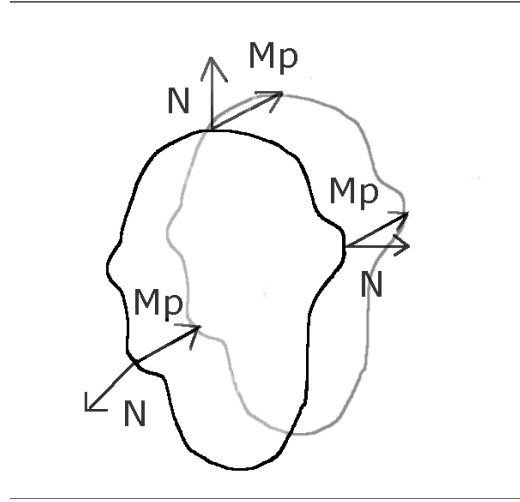


FIG. 8.2 – Différentes directions d'évolution : \mathbf{N} représente la normale de Γ_n et $M\mathbf{p}$ la direction du mouvement.

où $\mathbf{x} = \Gamma_n(p, \tau)$.

La particularité de cette équation d'évolution réside dans la direction d'évolution du contour actif.

Pour la segmentation d'images, le contour actif évoluait dans la direction de la normale.

Pour le suivi d'objets en mouvement, nous avons choisi de faire évoluer le contour dans la direction du mouvement estimé de l'objet.

8.3 Estimation du mouvement entre deux images

Une méthode classique d'estimation du mouvement entre deux images consécutives consiste à résoudre l'équation du flot optique. [2, 6, 39, 64, 78, 88]. Cependant, cette équation est sous-déterminée et nécessite des hypothèses supplémentaires quant au mouvement recherché. Même si des propriétés plus globales peuvent être prises en compte, à travers des termes de régularisation par exemple, cette approche est locale. En effet elle associe un vecteur à chaque pixel, alors que le mouvement dans une image est lié aux différents objets qui la compose, auquel s'ajoute éventuellement un mouvement de caméra global.

Enfin, l'équation du flot optique découle de l'approximation linéaire de

l'hypothèse d'intensité constante le long du déplacement, ce qui contraint l'estimation à de petits déplacements.

La méthode d'estimation de mouvement la plus répandue dans les normes de codage vidéo est la méthode de mise en correspondance de blocs ou *block-matching*. Cette méthode fait l'hypothèse d'un vecteur de mouvement par bloc. Elle permet d'avoir des vecteurs de mouvement plus cohérents avec les données, car un vecteur prend en compte l'information contenue dans tout le bloc qu'il représente. La limite de cette méthode vient du fait que la décomposition en blocs ne suit pas celle de l'image en objets. Le modèle de mouvement représentant un bloc est en général la translation. La méthode de recherche peut être exhaustive, pour assurer une meilleure estimation du mouvement, mais sera alors très coûteuse en temps de calcul. Certains algorithmes sous-optimaux [57, 92] permettent un gain de temps effectif, surtout pour une recherche sous-pixellique, mais au détriment de la qualité de l'estimation. Quelle que soit la méthode de recherche, l'amplitude du mouvement est limitée par une fenêtre de recherche, définie généralement à plus ou moins 15 pixels de déplacement, horizontalement et verticalement.

Notre méthode d'estimation de mouvement repose sur la minimisation d'une énergie respectant l'hypothèse d'intensité constante d'une région le long de son déplacement. Cependant, nous ne faisons pas d'approximation linéaire comme pour le flot optique, mais utilisons un algorithme de descente de gradient. De plus, nous supposons que le mouvement est représenté par un modèle paramétrique, constant sur une région, ce qui découle naturellement du contexte de segmentation de régions en mouvement dans lequel nous nous situons.

Nous écrivons le modèle de mouvement de la région Ω_n sous la forme $M(\mathbf{x}, n) \mathbf{p}$, où le vecteur \mathbf{p} contient les paramètres du modèle de mouvement. Les modèles de mouvement qui peuvent s'écrire sous la forme $M(\mathbf{x}, n) \mathbf{p}$ sont les modèles polynômiaux

$$\sum_{(i,j)/i+j \leq m} a_{i,j} x^i y^j$$

où m représente le degré du polynôme. En pratique les modèles les plus utilisés sont les modèles affines ($m = 1$), ou quadratiques ($m = 2$), car plus le degré du polynôme est élevé, plus le modèle est sensible au bruit.

Nous déduisons le modèle de mouvement de la minimisation par rapport à

\mathbf{p} , et pour Ω_n constant, de l'énergie suivante :

$$E(\mathbf{p}) = \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M(\mathbf{x}, n) \mathbf{p}))^2 d\mathbf{x} \quad (8.30)$$

En théorie, le modèle de mouvement $M(\mathbf{x}, n) \mathbf{p}$ peut représenter n'importe quel mouvement polynômial. En pratique, nous étudierons plus particulièrement les mouvements affines de type translation uniforme et homothétie.

8.3.1 Modèle de translation uniforme

Le calcul suivant s'applique dans le cas où la matrice $M(\mathbf{x}, n)$ ne dépend pas du point \mathbf{x} , c'est-à-dire $M(\mathbf{x}, n) = M(n)$.

L'exemple le plus simple est celui de la translation uniforme :

$$M(n) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{p} = \mathbf{a} \quad (8.31)$$

où \mathbf{a} est le vecteur de translation.

La minimisation du critère (8.30) passe par le calcul de sa dérivée par rapport à \mathbf{p} :

$$dE(\mathbf{p}) = 2 \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M(n)\mathbf{p})) M^T(n) \nabla I_{n-1}(\mathbf{x} + M(n)\mathbf{p}) d\mathbf{x} \quad (8.32)$$

où M^T est la matrice transposée de la matrice M .

Dans le cas de la translation, la dérivée du critère devient :

$$dE(\mathbf{p}) = 2 \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + \mathbf{a})) \nabla I_{n-1}(\mathbf{x} + \mathbf{a}) d\mathbf{x} \quad (8.33)$$

Les paramètres de mouvement $\mathbf{p}(\Omega_n)$ sont définis comme minima de l'énergie (8.30). Or cette énergie n'est pas forcément globalement convexe. De ce fait, nous ne pouvons utiliser un algorithme de descente de gradient, qui consiste à minimiser le critère linéairement dans les directions successives du gradient (ou du gradient conjugué). Nous pouvons toutefois considérer ce critère localement convexe autour de $\mathbf{p}(\Omega_n)$. Sous cette hypothèse, nous avons implémenté un algorithme de descente de gradient à pas variable.

8.3.2 Modèle de translation et d'homothétie centrée

Penchons-nous sur le cas où le modèle de mouvement comprend une homothétie et une translation. L'homothétie se définit à partir d'un centre \mathbf{g} (un point de D) et d'un facteur d'échelle k .

Il y a plusieurs façons de formuler la transformation comprenant une homothétie et une translation. Toutes sont en théorie équivalentes, mais ne donnent pas numériquement les mêmes résultats.

Soit \mathbf{q} le transformé du point \mathbf{x} après une homothétie puis une translation. Après l'homothétie on a :

$$\mathbf{q} - \mathbf{g} = k (\mathbf{x} - \mathbf{g}) \quad (8.34)$$

$$\iff \mathbf{q} = \mathbf{x} + (k - 1) \mathbf{x} - (k - 1) \mathbf{g} . \quad (8.35)$$

Puis \mathbf{q} est translaté d'un vecteur \mathbf{a}

$$\mathbf{q} = \mathbf{x} + (k - 1) \mathbf{x} - (k - 1) \mathbf{g} + \mathbf{a} \quad (8.36)$$

ce qui peut s'écrire :

$$\mathbf{q} = \mathbf{x} + \underbrace{\begin{pmatrix} -1 & 0 & x(\mathbf{x}) & 1 & 0 \\ 0 & -1 & y(\mathbf{x}) & 0 & 1 \end{pmatrix}}_{M(\mathbf{x},n)} \underbrace{\begin{pmatrix} (k-1) \mathbf{g} \\ k-1 \\ \mathbf{a} \end{pmatrix}}_{\mathbf{p}} \quad (8.37)$$

où $x(\mathbf{x})$ est l'abscisse du point \mathbf{x} , et $y(\mathbf{x})$ son ordonnée.

Remarquons que cela n'est pas équivalent à effectuer d'abord la translation puis l'homothétie. En effet, dans ce cas, le vecteur de translation serait multiplié par le facteur d'échelle k :

$$\mathbf{q} = \mathbf{x} + (k - 1) \mathbf{x} - (k - 1) \mathbf{g} + k \mathbf{a} \quad (8.38)$$

Cependant, résoudre le problème en estimant \mathbf{g} , k et \mathbf{a} est équivalent à résoudre le problème en estimant \mathbf{g} , k , et $k \mathbf{a}$.

Posons $\mathbf{b} = \mathbf{a} - (k - 1) \mathbf{g}$. L'équation (8.36) s'écrit :

$$\mathbf{q} = \mathbf{x} + (k - 1) \mathbf{x} + \mathbf{b} \quad (8.39)$$

ce qui peut s'interpréter comme une homothétie de facteur d'échelle k et de centre l'origine, suivi d'une translation de vecteur \mathbf{b} . De fait, on en

déduit que considérer l'origine de l'image comme centre de l'homothétie conserve toute la généralité du problème, et permet de n'avoir à estimer que 3 paramètres pour 2 inconnues (k, \mathbf{b}) , au lieu de 5 paramètres pour 3 inconnues $(\mathbf{g}, k, \mathbf{a})$ en utilisant les équations (8.36) et (8.38).

Nous allons choisir l'origine des coordonnées au barycentre \mathbf{g}_{Ω_n} de la région étudiée, plutôt qu'à l'origine de l'image, car ce choix induit moins d'erreurs. Le modèle de mouvement s'écrit alors :

$$M(\mathbf{x}, n) = \begin{pmatrix} x(\mathbf{x}) - x(\mathbf{g}_{\Omega_n}) & 1 & 0 \\ y(\mathbf{x}) - y(\mathbf{g}_{\Omega_n}) & 0 & 1 \end{pmatrix} \quad (8.40)$$

et

$$\mathbf{p} = \begin{pmatrix} k - 1 \\ \mathbf{b} \end{pmatrix} \quad (8.41)$$

où \mathbf{g}_{Ω_n} est le barycentre de Ω_n .

L'expression de la dérivée est maintenant directe puisqu'il suffit de remplacer $M(\mathbf{x}, n)$ et \mathbf{p} par leur expression (8.40) et (8.41) dans l'équation suivante :

$$dE(\mathbf{p}) = 2 \int_{\Omega_n} (I_n(\mathbf{x}_g) - I_{n-1}(k \mathbf{x}_g + b)) \quad (8.42)$$

$$(\mathbf{x}_g \text{ } Id_2)^T \quad \nabla I_{n-1}(k \mathbf{x}_g + b) \, d\mathbf{x}$$

où le vecteur $\mathbf{x}_g = \mathbf{x} - \mathbf{g}_{\Omega_n}$ est le vecteur des coordonnées de \mathbf{x} dans le repère d'origine le centre de gravité \mathbf{g}_{Ω_n} de la région Ω et Id_2 représente la matrice identité $2 * 2$.

Segmentation et estimation conjointes du mouvement sur un groupe d'images

Le mouvement d'un objet peut difficilement être défini de façon précise et complexe à partir de deux images. Un critère de segmentation et de suivi d'objets en mouvement a donc été défini sur un groupe d'images. Le critère se déduit directement du critère (8.4) sur deux images, en choisissant la dernière image de la séquence comme image de référence pour tout le groupe.

9.1 Définition du critère conjoint sur un groupe d'images

Le contour de l'objet d'intérêt évolue d'une image à l'autre en fonction de son mouvement. Or le mouvement d'un objet peut être complexe, et difficilement définissable à partir de deux images. Par exemple, on ne peut pas faire la différence entre une translation et une accélération si l'on ne prend en compte que les informations contenues dans deux images.

Lors de la définition de notre critère sur deux images (8.4) page 90, nous avons pris en compte un voisinage spatial afin d'accroître les informations disponibles et ainsi améliorer l'estimation du mouvement. De même, il semble judicieux, dans cette optique, d'utiliser les images voisines afin d'affiner notre estimation.

Le critère se définit donc temporellement sur un groupe d'images, et non plus sur deux images successives. Une extension directe du critère sur deux images consiste à propager le modèle de mouvement de proche en proche, en

sommant, pour chaque image du groupe considéré, le critère (8.4) entre cette image et sa précédente. Le critère serait alors le suivant :

$$E(\Omega_n) = \sum_{n=1}^N \int_{\Omega_n} (I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M(\mathbf{x}, n)\mathbf{p}(\Omega_n)))^2 d\mathbf{x} \quad (9.1)$$

Le vecteur $M(\mathbf{x}, n)\mathbf{p}(\Omega_n)$ représente le déplacement du point \mathbf{x} entre l'image I_n et l'image I_{n-1} . Le vecteur $\mathbf{p}(\Omega_n)$ contient les paramètres du modèle de mouvement défini sur la région Ω_n .

Cependant, nous avons préféré définir une **image de référence** pour le groupe, et calculer le mouvement par rapport à cette image de référence. En définissant cette image de référence, nous augmentons la robustesse de la segmentation, car nous évitons l'accumulation des erreurs que produit le critère qui prend en compte les images deux à deux successivement. En choisissant la dernière image du groupe comme image de référence, nous estimons le mouvement de l'image en cours à partir de la connaissance des images antérieures.

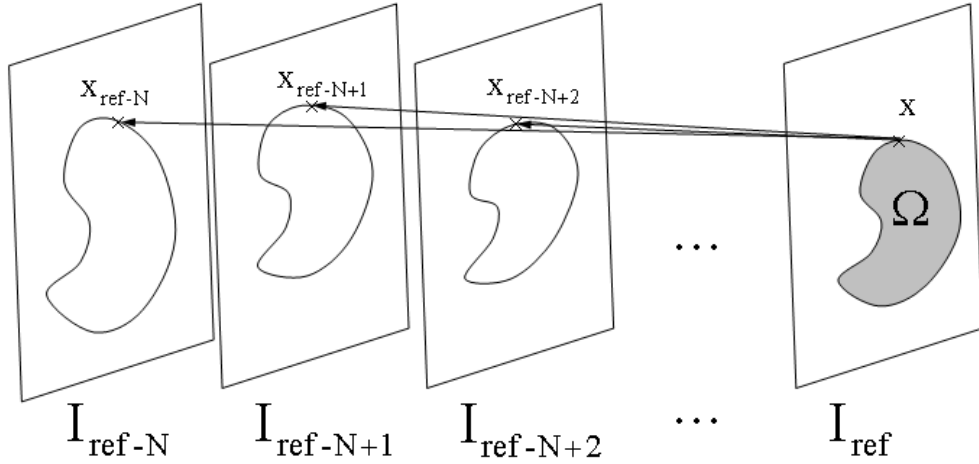


FIG. 9.1 – Modèle de mouvement sur un groupe d'images.

Nous proposons donc l'énergie suivante :

$$E(\Omega_{ref}) = \sum_{n=ref-1}^{ref-N} \int_{\Omega_{ref}} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega_{ref})))^2 d\mathbf{x} \quad (9.2)$$

La région Ω_{ref} représente l'objet dans l'image de référence et $\mathbf{x}_n(\Omega_{ref})$ est l'image du point \mathbf{x} par la transformation de paramètres $\mathbf{p}(\Omega_{ref})$ répétée $ref - n$ fois (voir figure 9.1). Par exemple, si le modèle de mouvement choisi est la translation : $\mathbf{x}_n(\Omega_{ref}) = \mathbf{x} + (ref - n) \mathbf{p}(\Omega_{ref})$.

L'estimation du modèle de mouvement est détaillée à la section 9.3.

Afin d'alléger les expressions des différentes équations, la région de référence Ω_{ref} sera notée plus simplement Ω , comme sur la figure 9.1.

Le vecteur des paramètres de mouvement $\mathbf{p}(\Omega)$ vérifie :

$$\mathbf{p}(\Omega) = \operatorname{argmin}_{\mathbf{p}} \sum_{n=ref-1}^{ref-N} \int_{\Omega} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega)))^2 d\mathbf{x} . \quad (9.3)$$

9.2 Segmentation et suivi d'objets en mouvement

Pour segmenter l'image I_{ref} en mouvement, nous minimisons l'énergie (9.2) par rapport à la région Ω . Nous calculons donc sa dérivée par rapport à Ω .

Une famille de transformations est définie de façon à rendre le domaine Ω dépendant d'un schéma dynamique τ .

Corollaire 9.1. *La dérivée Eulérienne dans la direction \mathbf{V} du critère :*

$$E(\Omega) = \sum_{n=ref-1}^{ref-N} \int_{\Omega} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega)))^2 d\mathbf{x} \quad (9.4)$$

où les paramètres du modèle de mouvement $\mathbf{p}(\Omega)$

$$\mathbf{p}(\Omega) = \min_{\mathbf{p}} \sum_{n=ref-1}^{ref-N} \int_{\Omega} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega)))^2 d\mathbf{x} . \quad (9.5)$$

a pour expression

$$dE(\Omega, \mathbf{V}) = - \sum_{n=ref-1}^{ref-N} \int_{\Gamma} (I_{ref}(s) - I_n(s_n(\Omega)))^2 \mathbf{V} \cdot \mathbf{N} ds \quad (9.6)$$

où Γ représente le contour orienté de Ω , s est l'abscisse curviligne du point $\mathbf{x} = \Gamma(s)$, $s_n(\Omega)$ est l'abscisse curviligne du point $\mathbf{x}_n = \Gamma_n(s_n(\Omega))$ image de \mathbf{x} par le modèle de mouvement $\mathbf{p}(\Omega)$ composé $ref - n$ fois, \mathbf{V} est le vecteur de déformation (ou vecteur de vitesse), inconnu, de Γ , et \mathbf{N} la normale unitaire intérieure de Γ .

Démonstration :

La dérivée est calculée dans la direction \mathbf{V} , par la méthode des gradients de forme [26, 51] :

$$\begin{aligned} dE(\Omega, \mathbf{V}) &= \sum_{n=ref-1}^{ref-N} \int_{\Omega} d_{\Omega}((I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega)))^2)(\mathbf{x}, \Omega, \mathbf{V}) \, d\mathbf{x} \\ &\quad - \sum_{n=ref-1}^{ref-N} \int_{\Gamma} (I_{ref}(s) - I_n(s_n(\Omega)))^2 \mathbf{V} \cdot \mathbf{N} \, ds \end{aligned} \quad (9.7)$$

Dans la démonstration du théorème 8.8 page 92, nous avons vu que

$$\int_{\Omega_n} d_{\Omega}((I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + M \mathbf{p}(\Omega)))^2)(\mathbf{x}, \Omega, \mathbf{V}) \, d\mathbf{x} = 0 \quad (9.8)$$

En suivant le même raisonnement, on conclut :

$$\int_{\Omega} d_{\Omega}((I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega)))^2)(\mathbf{x}, \Omega, \mathbf{V}) \, d\mathbf{x} = 0, \quad \forall n, ref-N \leq n \leq ref-1 \quad (9.9)$$

Finalement, la dérivée du critère sur un groupe d'images est :

$$dE(\Omega, \mathbf{V}) = - \sum_{n=ref-1}^{ref-N} \int_{\Gamma} (I_{ref}(s) - I_n(s_n(\Omega)))^2 \mathbf{V} \mathbf{N} \, ds \quad (9.10)$$

□

Le principe des contours actifs consiste à faire évoluer un contour initial vers l'objet suivant une force, dont la direction est donnée, en général par le vecteur normal au contour, \mathbf{N} . Cette force est calculée de façon à minimiser le critère d'énergie décrivant l'objet. On s'assure que le contour se dirige vers le minimum du critère en choisissant la force d'évolution du contour telle que la dérivée du critère soit négative.

9.2.1 Équation d'évolution pour la segmentation d'objets en mouvement

Dans le cadre d'une application de type segmentation, la direction d'évolution du contour choisie est celle de la normale au contour, \mathbf{N} .

L'équation d'évolution déduite de la dérivée du critère (9.10) est :

$$\frac{\partial \Gamma}{\partial \tau}(p, \tau) = \alpha \sum_{n=ref-1}^{ref-N} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega(\tau))))^2 \mathbf{N} \quad (9.11)$$

où α est une constante positive qui peut être considérée comme un pas d'évolution, par analogie avec le pas de descente d'un algorithme de descente de gradient. On note $\mathbf{x} = \Gamma(p, \tau)$ et $\mathbf{x}_n(\Omega(\tau))$ l'image de \mathbf{x} par le modèle de mouvement $\mathbf{p}(\Omega(\tau))$ composé $ref - n$ fois.

Remarquons, de plus, que l'évolution du contour nécessite l'évaluation des paramètres de mouvement à chaque itération (voir section 9.3).

9.2.2 Équation d'évolution pour le suivi d'objets en mouvement

Dans le cadre du suivi d'objets en mouvement, et pour les raisons évoquées à la section 8.2.2 page 96, la direction d'évolution du contour actif suit celle du mouvement de l'objet. L'équation d'évolution du contour actif se présente comme suit :

$$\frac{\partial \Gamma}{\partial \tau}(p, \tau) = \alpha \sum_{n=ref-1}^{ref-N} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega(\tau))))^2 M \mathbf{p}(\Omega(\tau)) \quad (9.12)$$

L'utilisation des paramètres de mouvement permet, lors de l'évolution du contour actif, de prendre en compte des propriétés globales à l'objet, et non plus seulement locales comme pour la segmentation.

9.3 Estimation du mouvement

Le vecteur des paramètres de mouvement $\mathbf{p}(\Omega)$ vérifie :

$$\mathbf{p}(\Omega) = \operatorname{argmin}_{\mathbf{p}} \sum_{n=ref-1}^{ref-N} \int_{\Omega} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega)))^2 d\mathbf{x} . \quad (9.13)$$

Le vecteur des paramètres de mouvement $\mathbf{p}(\Omega)$ reste constant sur la région qu'il décrit tout au long de la séquence de $N + 1$ images. En d'autres termes, on peut estimer le point \mathbf{x}_{n-1} de l'image I_{n-1} correspondant au point \mathbf{x}_n de l'image I_n à partir de l'équation :

$$\mathbf{x}_{n-1} = \mathbf{x}_n + M(\mathbf{x}_n, n) \mathbf{p}(\Omega) , \quad \forall n \text{ tel que } ref - 1 \leq n \leq ref - N \quad (9.14)$$

Cependant, le critère est défini entre un point \mathbf{x} de l'image de référence I_{ref} et son correspondant \mathbf{x}_n dans l'image I_n , pour n compris entre $ref - 1$ et $ref - N$. On doit donc évaluer l'expression du point \mathbf{x}_n par rapport au point de référence \mathbf{x} . De proche en proche on obtient :

$$\begin{aligned} \mathbf{x}_{ref-1} &= \mathbf{x} + M(\mathbf{x}, n_{ref}) \mathbf{p}(\Omega) \\ \mathbf{x}_{ref-2} &= \mathbf{x}_{ref-1} + M(\mathbf{x}_{ref-1}, n_{ref-1}) \mathbf{p}(\Omega) \\ &= \mathbf{x} + [M(\mathbf{x}, n_{ref}) + M(\mathbf{x}_{ref-1}, n_{ref-1})] \mathbf{p}(\Omega) \\ \mathbf{x}_{ref-3} &= \mathbf{x} + [M(\mathbf{x}, n_{ref}) + M(\mathbf{x}_{ref-1}, n_{ref-1}) + M(\mathbf{x}_{ref-2}, n_{ref-2})] \mathbf{p}(\Omega) \end{aligned}$$

Soit pour un point \mathbf{x}_n de l'image I_n :

$$\mathbf{x}_n = \mathbf{x} + \underbrace{[M(\mathbf{x}, n_{ref}) + \cdots + M(\mathbf{x}_{ref-n+1}, n_{ref-n+1})]}_{ref-n \text{ termes}} \mathbf{p}(\Omega) \quad (9.15)$$

L'expression de \mathbf{x}_n dans l'équation (9.15) fait intervenir les matrices $M(\mathbf{x}, n_{ref})$ à $M(\mathbf{x}_{n+1}, n_{ref-n+1})$, ce qui nécessite sous cette forme de nombreux calculs numériques. Cependant, pour des modèles de mouvement de type translation uniforme et homothétie, l'expression de \mathbf{x}_n en fonction de \mathbf{x} s'avère beaucoup plus simple.

9.3.1 Modèle de translation uniforme

Dans le cas où la matrice $M(\mathbf{x}, n)$ est indépendante de \mathbf{x} et de n , soit $M(\mathbf{x}, n) = M$, l'expression de \mathbf{x}_n en fonction de \mathbf{x} , point de la région Ω dans l'image I_{ref} , est simple :

$$\mathbf{x}_n(\Omega) = \mathbf{x} + (ref - n) M \mathbf{p}(\Omega) \quad \forall n \text{ tel que } ref - 1 \leq n \leq ref - N \quad (9.16)$$

Par exemple, pour une translation uniforme, $M(\mathbf{x}, n) = Id$, où Id est la matrice identité :

$$\mathbf{x}_n(\Omega) = \mathbf{x} + (ref - n) \mathbf{p}(\Omega) \quad \forall n \text{ tel que } ref - 1 \leq n \leq ref - N \quad (9.17)$$

La dérivée de l'énergie (9.13) par rapport aux paramètres de mouvement \mathbf{p} vaut :

$$dE(\mathbf{p}) = -2 \sum_{n=ref-1}^{ref-N} \int_{\Omega} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x} + (ref - n) M\mathbf{p})) (ref - n) M^T \nabla I_n(\mathbf{x} + (ref - n) M\mathbf{p}) d\mathbf{x} . \quad (9.18)$$

où M est la matrice identité dans le cas d'une translation uniforme.

9.3.2 Modèle de translation uniforme et d'homothétie centrée

Pour un modèle de mouvement de type homothétie centrée suivi d'une translation, si l'on garde la notation

$$\mathbf{x}_{n-1} = \mathbf{x}_n + M(\mathbf{x}_n, n) \mathbf{p}(\Omega) , \quad \forall n \text{ tel que } ref - 1 \leq n \leq ref - N \quad (9.19)$$

on a :

$$M(\mathbf{x}, n) = \begin{pmatrix} x(\mathbf{x}) - x(\mathbf{g}_{\Omega}) & 1 & 0 \\ y(\mathbf{x}) - y(\mathbf{g}_{\Omega}) & 0 & 1 \end{pmatrix} \text{ et } \mathbf{p} = \begin{pmatrix} k - 1 \\ \mathbf{a} \end{pmatrix} \quad (9.20)$$

où \mathbf{g}_{Ω} est le barycentre de Ω . L'expression de \mathbf{x}_n n'est alors guère plus simple que celle de l'équation (9.15).

Afin d'éviter l'estimation successive des matrices $M(\mathbf{x}, n_{ref})$ à $M(\mathbf{x}_{n+1}, n_{ref-n+1})$ (voir équation (9.15)), nous avons préféré exprimer \mathbf{x}_{n-1} comme l'image de \mathbf{x}_n par l'homothétie de centre \mathbf{g}_{Ω} , barycentre de la région Ω , et de facteur d'échelle k puis la translation de vecteur \mathbf{b} , c'est-à-dire comme suit :

$$\mathbf{x}_n - \mathbf{g}_{\Omega} = k(\mathbf{x}_{n-1} - \mathbf{g}_{\Omega}) + \mathbf{b} \quad (9.21)$$

Notons $\mathbf{x}_{n,\mathbf{g}} = \mathbf{x}_n - \mathbf{g}_{\Omega}$ les coordonnées de \mathbf{x}_n dans le repère d'origine \mathbf{g}_{Ω} , l'équation précédente s'écrit alors :

$$\mathbf{x}_{n,\mathbf{g}} = k \mathbf{x}_{n-1,\mathbf{g}} + \mathbf{b} \quad (9.22)$$

L'expression du point \mathbf{x}_n s'écrit alors simplement :

$$\mathbf{x}_{n,\mathbf{g}} = k^{ref-n} \mathbf{x}_{\mathbf{g}} + \mathbf{b} \sum_{j=0}^{ref-n-1} k^j \quad (9.23)$$

où $\mathbf{x}_{\mathbf{g}}$ représente les coordonnées du point \mathbf{x} dans le repère d'origine \mathbf{g}_{Ω} .

La dérivée du critère d'énergie (9.13) par rapport aux paramètres de mouvement \mathbf{p} vaut :

$$dE(\mathbf{p}) = -2 \sum_{n=ref-1}^{ref-N} \int_{\Omega} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega))) \left((ref - n)k^{ref-n-1} \mathbf{x} + \mathbf{b} \sum_{j=0}^{ref-n-1} j k^{j-1} \sum_{j=0}^{ref-n-1} k^j \right)^T \nabla I_n(\mathbf{x}_n) d\mathbf{x} \quad (9.24)$$

Chapitre 10

Résultats expérimentaux

Dans ce chapitre nous allons appliquer le cadre de travail précédent à la segmentation et au suivi d'objets en mouvement.

10.1 Segmentation/estimation conjointes du mouvement

Nous évaluerons tout d'abord l'estimation du mouvement en la comparant à l'estimation par *block-matching* puis par des méthodes classiques d'estimation du flot optique. Ensuite nous validerons la segmentation d'images réelles.

10.1.1 Validation de l'estimation du mouvement

Afin d'évaluer la qualité de l'estimation du mouvement, nous allons définir une région sur laquelle nous estimerons le mouvement suivant l'équation (voir section 9.3) :

$$\mathbf{p}(\Omega) = \operatorname{argmin}_{\mathbf{p}} \sum_{n=\operatorname{ref}-1}^{\operatorname{ref}-N} \int_{\Omega} (I_{\operatorname{ref}}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega)))^2 d\mathbf{x} . \quad (10.1)$$

Nous comparerons notre estimation à deux types de méthodes : le *block-matching* et des méthodes classiques d'estimation du flot optique.

Comparaison avec une estimation de type *block-matching*

Pour cette première comparaison, la région d'intégration Ω est un bloc. Pour se placer dans des conditions équivalentes à celles de l'estimation par *block-matching*, le modèle de mouvement représente une translation entre

deux images (voir section 8.3.1 page 99).

Les paramètres de l'estimation par *block-matching* sont classiques : le critère de comparaison du *block-matching* est la somme des différences au carré (SSD), la méthode de recherche est exhaustive, la fenêtre de recherche va de -15 pixels à $+15$ pixels horizontalement et verticalement, et la précision est au demi-pixel.

Contrairement au *block-matching*, notre méthode n'impose aucune limite d'amplitude ou de précision.

Les résultats obtenus pour la séquence Foreman sont présentés en figure 10.1.

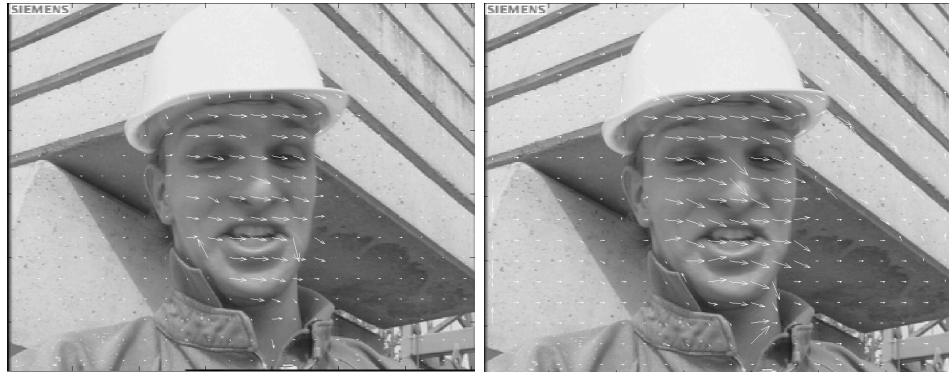


FIG. 10.1 – Estimation du mouvement entre les images 80 et 79 de la séquence Foreman : à gauche, l'estimation par l'équation (10.1), à droite, l'estimation par *block-matching*.

Les critères de comparaison entre deux blocs, aussi bien pour notre méthode que pour le *block-matching*, perdent leur sens sur des régions homogènes. Dans les deux cas, tout bloc dont la variance est inférieure à un certain seuil est considéré comme homogène et se voit assigner un vecteur de mouvement nul (on retrouve de tels blocs sur le casque de Foreman).

Pour évaluer quantitativement les résultats, nous estimons l'image 80 par une prédiction *backward* pour chaque estimation des vecteurs de mouvement. La prédiction d'une image en *backward* suit le même principe que la compensation du mouvement de la caméra présentée en Annexe A.2 page 142. De l'image des erreurs entre l'image 80 réelle et celle prédite, nous déduisons le rapport signal à bruit (PSNR : peak signal to noise) de chacune des méthodes. La méthode proposée améliore le PSNR de près de 1dB, son PSNR est de 33.87dB contre 32.88 pour le *block-matching*.

Les histogrammes de l'erreur de prédiction pour chaque méthode confirment ces résultats (voir figure 10.2).

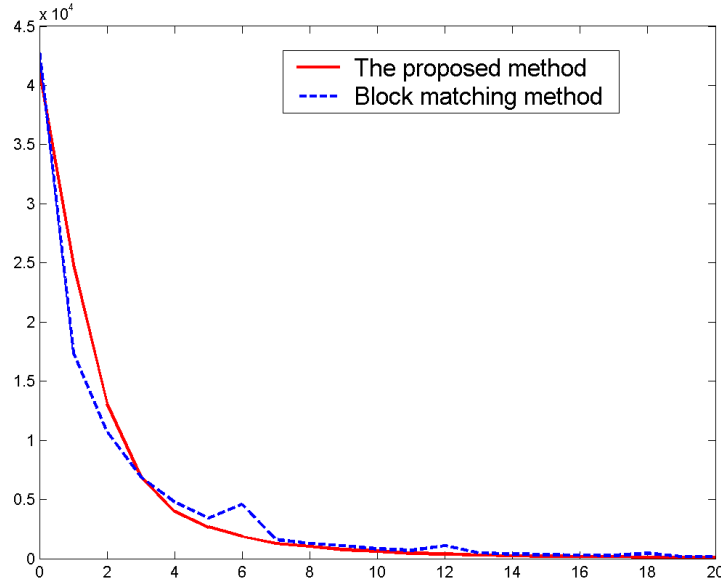


FIG. 10.2 – Histogrammes de l'erreur de prédiction pour l'image 80 : en rouge notre méthode, en bleu le *block-matching*.

Si l'estimation par la résolution de l'équation (10.1) donne, dans ce cas, de meilleurs résultats que le *block-matching*, elle a toutefois l'inconvénient de dépendre de son initialisation. Le risque est de trouver un minimum local, alors que le *block-matching* assure, pour une précision et une fenêtre de recherche fixées, de trouver le minimum global. Une solution qui allie les avantages des deux méthodes consiste à initialiser l'estimation de l'équation (10.1) par un *block-matching* exhaustif au pixel, par exemple. Cette méthode sera mise en oeuvre pour la segmentation et le suivi d'objets en mouvement, présentés aux sections suivantes, mais ne sera pas utilisée pour la comparaison avec les méthodes classiques de flot optique.

Comparaison avec des méthodes classiques de flot optique

L'évaluation de nos résultats par rapport à des méthodes de flot optique s'appuie sur les travaux de Barron *et al.* [4]. Nous comparons nos résultats, obtenus sur une de leurs séquences de test, *translating tree*, dont le mouvement réel est connu, avec ceux obtenus par des méthodes d'estimation du flot optique, suivant des critères définis par Barron *et al.*. Le mouvement, connu, de la séquence, est un mouvement de translation

horizontale, dont l'amplitude va croissante de la gauche à la droite de l'image.

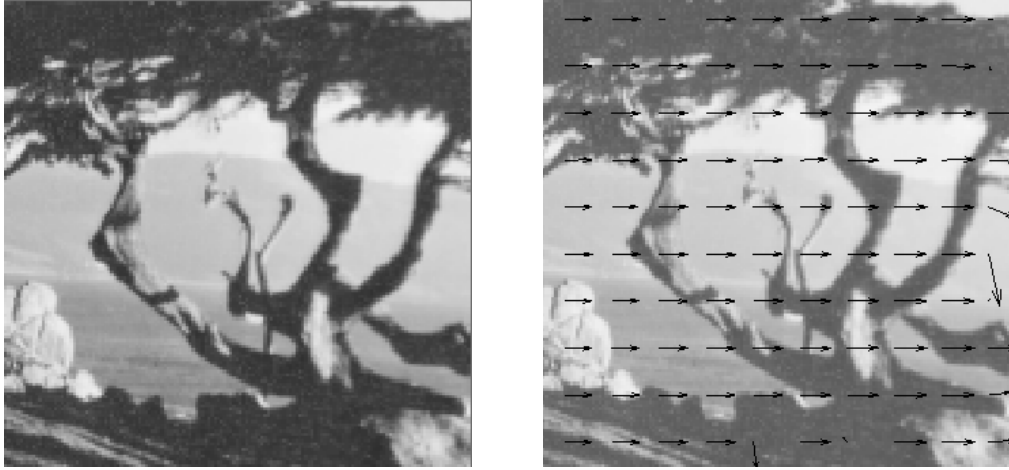


FIG. 10.3 – A gauche : l'image originale ; A droite : les vecteurs de déplacement obtenus en estimant le mouvement par blocs 15×15 sur 3 images.

Notre méthode s'appuie sur l'estimation du mouvement sur une région. Nous découpons l'image en blocs 15×15 (l'image a une taille de 150×150 pixels) et estimons le mouvement sur chacun de ces blocs. Nous assignons à chaque pixel du bloc le mouvement du bloc, afin de comparer notre méthode à des méthodes estimant des champs denses. Cette méthode nous désavantage légèrement car le mouvement réel est différent pour chaque pixel.

L'estimation du mouvement est effectuée entre 3 images, et pour un modèle de mouvement de translation (voir figure 10.3).

TAB. 10.1 – Comparaison de l'estimation de mouvement avec les résultats de Barron *et al.* [4].

Technique	ϵ	σ	δ
Horn et Schunck (original)	38.72°	27.67°	100%
Uras <i>et al.</i> (sans seuil)	0.62°	0.52°	100%
Fleet et Jepson ($\tau = 1.25$)	0.23°	0.19°	49.7%
Méthode proposée	0.33°	0.21°	81%

Le tableau 10.1 présente des résultats extraits de [4], plus précisément ceux de la meilleure méthode dont la densité est supérieure à 75%, ceux de la meilleure méthode toutes densités confondues, et les résultats de la

méthode classique de Horn et Schunck. La ligne en gras présente nos résultats.

Notations : ϵ est l'erreur angulaire moyenne, σ la déviation standard de l'erreur, et δ est la densité des pixels pris en compte dans le calcul de ϵ et σ .

Estimation du mouvement sur une région

La figure 10.4 présente un résultat obtenu sur la séquence *Foreman* avec l'énergie calculée sur 3 images et un modèle de mouvement de translation. Cette fois la région Ω a été segmentée manuellement et détoure le visage et les épaules de *Foreman*. Une estimation par *block-matching* exhaustif, au

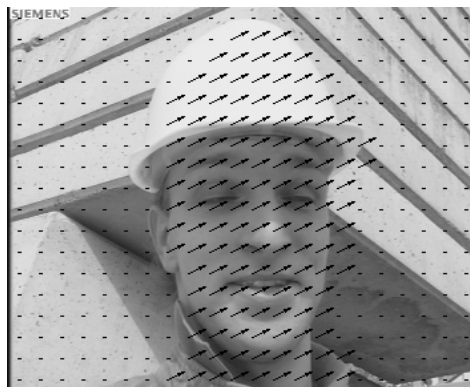


FIG. 10.4 – Estimation du mouvement par région sur 3 images de la séquence *Foreman*.

pixel et sur des blocs 16×16 , sert d'initialisation à l'algorithme de descente de gradient. Pour une meilleure visualisation des vecteurs de mouvement, nous représentons un vecteur de mouvement par bloc au lieu d'un seul vecteur pour toute la région.

10.1.2 Validation de la segmentation conjointe

Nous appliquons l'algorithme d'estimation et de segmentation conjointes du mouvement sur 2 images consécutives de la séquence *Coastguard*. Le critère est défini sur un groupe de trois images pour un modèle de mouvement de translation.

Les résultats sont présentés à la figure 10.5.

Remarquez que l'antenne à l'avant du bateau (presque invisible sur les deux images) arrête l'évolution du contour actif de l'image 251, l'empêchant de

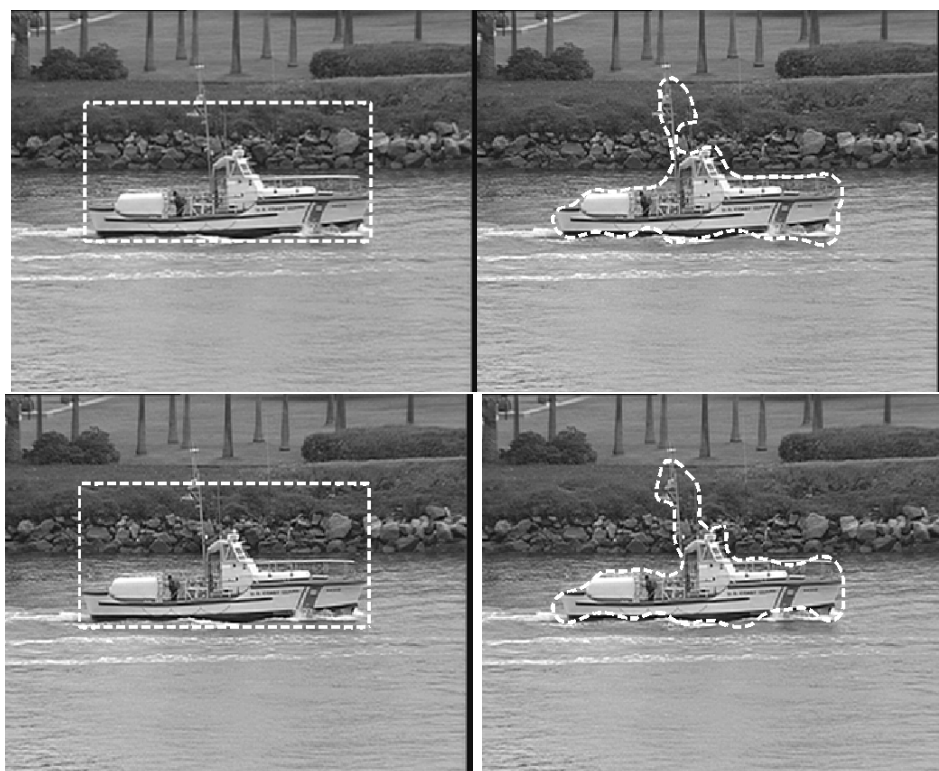


FIG. 10.5 – Estimation et segmentation conjointes du mouvement de 2 images successives de la séquence *Coastguard* sur un groupe de 3 images. Sur la première ligne : le contour initial et la segmentation de l'image 250 ; Sur la deuxième ligne : *Idem* pour l'image 251.

s'approcher de l'antenne principale du bateau. Ce phénomène s'observe aussi sur l'image 250 mais est moins prononcé et ne concerne que le haut de l'antenne.

10.2 Suivi d'objets en mouvement

10.2.1 Directions d'évolution du contour actif

Nous créons une séquence synthétique en superposant le visage de Foreman à un fond d'image. Le fond restant fixe, nous appliquons au visage de Foreman une translation horizontale de 3 pixels entre chaque image de la séquence synthétique. La première image est segmentée manuellement.

Nous souhaitons comparer, pour les mêmes images, les directions d'évolution

du contour actif, suivant la normale, et suivant le mouvement.

Pour l'équation d'évolution du contour actif **suivant la normale**, le suivi d'objets en mouvement est mis en oeuvre par un algorithme de contours actifs, dont le critère est défini par une compétition de régions entre Ω^{out} le fond de l'image de référence, et Ω^{in} l'objet de référence :

$$\begin{aligned} E(\Omega^{in}, \Omega^{out}) = & \sum_{n=ref-1}^{ref-N} \left(\int_{\Omega^{in}} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{in})))^2 d\mathbf{x} \right. \\ & \left. + \int_{\Omega^{out}} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{out})))^2 d\mathbf{x} \right) \end{aligned} \quad (10.2)$$

où

$$\mathbf{p}(\Omega^{in}) = \operatorname{argmin}_{\mathbf{p}} \sum_{n=ref-1}^{ref-N} \int_{\Omega^{in}} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{in})))^2 d\mathbf{x} \quad (10.3)$$

et

$$\mathbf{p}(\Omega^{out}) = \operatorname{argmin}_{\mathbf{p}} \sum_{n=ref-1}^{ref-N} \int_{\Omega^{out}} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{out})))^2 d\mathbf{x} \quad (10.4)$$

L'équation d'évolution du contour actif est dirigée suivant la normale :

$$\begin{aligned} \frac{\partial \Gamma}{\partial \tau}(p, \tau) = & \alpha \left(\sum_{n=ref-1}^{ref-N} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{in}(\tau))))^2 \right. \\ & \left. + \sum_{n=ref-1}^{ref-N} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{out}(\tau))))^2 \right) \mathbf{N} \end{aligned} \quad (10.5)$$

où le modèle de mouvement est une translation :

$\mathbf{x}_n = \mathbf{x} + (n - ref)\mathbf{p}(\Omega(\tau))$ avec $\mathbf{x} = \Gamma(p, \tau)$ pour $\Omega = \Omega^{in}$ ou Ω^{out} .

Pour une équation d'évolution dirigée suivant la normale, la compétition de régions est nécessaire si l'on souhaite que le contour puisse évoluer vers l'intérieur ou vers l'extérieur du contour. En effet si le critère n'était composé que du terme défini sur l'objet, pour minimiser le critère, le contour actif évoluerait de sorte à minimiser l'aire de l'objet. Le terme défini sur le fond de l'image est nécessaire pour compenser l'influence du terme défini sur l'objet, et permettre au contour d'évoluer dans les deux sens.

En conséquence, l'algorithme sera sensible aux occlusions du fond (voir le détail de la figure 10.6.a).

Pour une équation d'évolution **suivant le mouvement**, seuls les termes sur l'objets sont nécessaires. L'équation d'évolution est :

$$\frac{\partial \Gamma}{\partial \tau}(p, \tau) = \alpha \left(\sum_{n=ref-1}^{ref-N} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{in}(\tau))))^2 \right) M \mathbf{p}(\Omega^{in}(\tau)) \quad (10.6)$$

où le modèle de mouvement est une translation :

$$\mathbf{x}_n(\Omega^{in}(\tau)) = \mathbf{x} + (n - ref) \mathbf{p}(\Omega^{in}(\tau)) \text{ avec } \mathbf{x} = \Gamma(p, \tau).$$

L'objet est segmenté suivant son mouvement. Suivant l'évolution du contour, le mouvement de la région $\Omega^{in}(\tau)$ peut changer de sens, alors que la normale intérieure au contour ne définit qu'un seul sens de propagation. Lors d'une évolution suivant le mouvement, l'occlusion du fond de l'image n'a pas d'effet sur la qualité de la segmentation (voir le détail de la figure 10.6.b). Une occlusion de l'objet lui-même, par contre, ne serait pas sans conséquence.



a. direction d'évolution suivant la normale b. direction d'évolution suivant le mouvement

FIG. 10.6 – Suivi d'objet en mouvement sur la séquence synthétique créée à partir de Foreman.

10.2.2 Résultats en suivi d'objets en mouvement

Nous avons appliqué à la séquence *flower and garden*, l'équation d'évolution :

$$\frac{\partial \Gamma}{\partial \tau}(p, \tau) = \alpha \left(\sum_{n=ref-1}^{ref-N} (I_{ref}(\mathbf{x}) - I_n(\mathbf{x}_n(\Omega^{in}(\tau))))^2 \right) M \mathbf{p}(\Omega^{in}(\tau)) \quad (10.7)$$

où le modèle de mouvement est une translation :

$$\mathbf{x}_n(\Omega^{in}(\tau)) = \mathbf{x} + (n - ref)\mathbf{p}(\Omega^{in}(\tau)), \text{ avec } \mathbf{x} = \Gamma(p, \tau).$$

La première image a été segmentée manuellement, et le suivi a été effectué sur 15 images consécutives. La figure 10.7 présente quatre images de la séquence extraites toutes les 5 images. Le haut de l'arbre, plus précisément les



FIG. 10.7 – Suivi d'objets en mouvement sur 15 images de la séquence *flower and garden* avec l'équation d'évolution suivant la direction du mouvement.

branches, n'est pas très bien segmenté, cela est probablement dû au modèle de mouvement trop simple.

Quatrième partie

Conclusions et perspectives

Conclusions

Dans ce mémoire nous proposons des modèles de contours actifs pour la segmentation d'images et de vidéos.

Notre approche est variationnelle, c'est-à-dire que le problème de segmentation est formulé comme la minimisation d'un critère. La dérivation du critère permet de déduire l'équation d'évolution du contour actif.

Deux types de termes peuvent constituer un critère de segmentation : des termes de contour et des termes de région. Dans les deux cas, un terme s'écrit comme l'intégrale sur la région (ou le contour) d'une fonction appelée descripteur de la région (ou du contour). Certains descripteurs de région apportent des informations globales sur les propriétés des régions, comme leur moyenne, leur variance, leur entropie, leur histogramme ou leur mouvement. Les termes de région correspondants dépendent alors doublement de la région : comme intégrale sur la région d'un descripteur dépendant lui-même de la région. On appelle ces termes à *double dépendance*. En fonction de la dépendance du critère à la région, la dérivation du critère par rapport à la région n'est pas aisée. Nous utilisons à cette fin des outils de dérivation empruntés à l'optimisation de domaine : les gradients de forme.

Nous présentons, dans ce mémoire, la dérivation de chaque type de termes effectuée par la méthode des gradients de forme, en portant une attention particulière à la dérivation des termes de région à double dépendance.

La contribution de cette thèse réside dans l'élaboration et l'étude de deux descripteurs particuliers de régions.

Le premier descripteur définit un a priori géométrique sans contrainte paramétrique : l'objectif est de minimiser la distance euclidienne du contour actif à un contour de référence. Nous ne faisons pas l'hypothèse d'une transformation paramétrique entre le contour actif et le contour de référence, et notre critère ne dépend pas d'une paramétrisation particulière du contour.

En ce sens, notre critère n'impose aucune contrainte paramétrique au contour actif. Nous proposons deux dérivations possibles du critère d'a priori : par le calcul des variations, et par les gradients de forme, et en déduisons l'équation d'évolution du contour actif pour ce critère. Puis nous proposons trois applications du critère d'a priori géométrique.

Utilisé seul, ce critère est appliqué à la déformation de contours ou *shape warping*. Les résultats obtenus montrent qu'il n'y a aucune contrainte paramétrique de forme sur le contour actif. Le contour se rapproche du contour de référence en se déformant librement.

Pour des applications de segmentation de régions, nous combinons le terme d'a priori géométrique avec un terme de région statistique sur l'objet, et un terme de régularisation classique. Afin de montrer l'intérêt du terme d'a priori, nous comparons notre méthode à une compétition de régions utilisant le même descripteur statistique, sur le fond de l'image et sur l'objet. Le terme d'a priori géométrique nous dispense d'utiliser un terme de région décrivant le fond de l'image. En effet, cet a priori entre en compétition avec les informations statistiques de l'objet, et les informations statistiques du fond de l'image ne sont plus nécessaires. Le terme d'a priori permet à l'utilisateur d'interagir en définissant un contour de référence détournant grossièrement l'objet d'intérêt.

Le terme d'a priori géométrique, associé à un descripteur statistique de la région de l'objet et à un terme de régularisation, est appliqué au suivi de région. Le terme d'a priori permet d'introduire une cohérence temporelle. Cette cohérence est liée au fait que le contour de référence, défini une seule fois par l'utilisateur sur la première image, se déduit ensuite du contour final de l'image précédente, pour les autres images de la séquence.

Le deuxième descripteur définit un critère de segmentation d'image en objets en mouvement, conjointement à l'estimation du mouvement de la région segmentées. La plupart des méthodes d'estimation du mouvement font l'hypothèse, communément admise, que l'intensité lumineuse d'un pixel ne varie pas le long de sa trajectoire. Mais le problème ainsi posé est sous-déterminé. Nous posons l'hypothèse d'un mouvement constant par région, en étendant le domaine d'application de l'hypothèse, d'un point à une région. Nous restreignons également le mouvement en le représentant par un modèle de mouvement paramétrique.

Le critère est tout d'abord défini sur deux images successives, puis pour un groupe d'images, dont la dernière est choisie comme image de référence.

Nous dérivons le critère par rapport à la région par la méthode des gradients de forme, et proposons deux équations d'évolution du contour actif : une dirigée suivant la normale au contour, pour la segmentation d'objets en

mouvement, et une dirigée suivant le mouvement de l'objet, pour le suivi d'objets en mouvement.

Nous étudions la méthode d'estimation du mouvement définie conjointement par le critère, pour un modèle de mouvement polynômial, en nous attachant plus particulièrement aux modèles de translation uniforme et de d'homothétie.

Nous avons appliqué ce critère à l'estimation et la segmentation conjointe du mouvement. Nous évaluons la qualité de l'estimation du mouvement, pour une région donnée, en la comparant à des méthodes classiques d'estimation du flot optique et à une méthode de mise en correspondance, sur une séquence synthétique dont le mouvement est connu. Puis, nous appliquons le critère conjoint à la segmentation d'images réelles. Le critère de segmentation comprend alors le terme de segmentation et d'estimation conjointe défini sur l'objet, et le même terme défini sur le fond de l'image, établissant une compétition entre les deux régions. Le contour actif évolue suivant la direction de sa normale intérieure.

Nous présentons les résultats obtenus en suivi d'objets en mouvement en comparant les deux directions d'évolution du contour actif : suivant la normale au contour et suivant le mouvement. Lorsque l'équation d'évolution est dirigée suivant la normale, la compétition de régions est nécessaire si l'on souhaite que le contour puisse évoluer dans les deux sens de la direction. En conséquence, l'algorithme est sensible aux zones de recouvrement du fond. Lorsque l'équation d'évolution est dirigée suivant le mouvement, seul le terme sur l'objet est nécessaire. Les zones de recouvrement du fond de l'image n'a alors pas d'effet sur la qualité de la segmentation.

Perspectives

Les descripteurs définis dans ce document et les résultats obtenus permettent d'ouvrir de nouveaux axes de réflexions.

D'un point de vue applicatif, le critère d'a priori géométrique pourrait fournir un cadre de travail pour le développement d'une méthode de *morphing*, technique utilisée en post-production, ainsi que pour l'interpolation temporelle d'une séquence vidéo, nécessaire au codage prédictif et à la conversion de standards.

Le critère d'a priori géométrique définit une carte de distance du contour actif au contour de référence. Le choix d'une implémentation par ensembles de niveaux introduit le concept de carte des distances au contour actif. En posant l'hypothèse de conservation des ensembles de niveaux, on pourrait

spécifier une correspondance entre les surfaces définies par les contours actifs et de référence.

En prenant comme surface deux visages, cela reviendrait à effectuer le *morphing* d'un visage à l'autre. En prenant comme surface deux images successives d'une vidéo, cela permettrait de sur-échantillonner la vidéo.

En imagerie médicale, un contour de référence est généralement défini à partir d'un atlas ou d'un modèle statistique, créé par apprentissage à partir d'un certain nombre de courbes. Ces applications nécessitent l'introduction d'un a priori plus contraint que celui défini par notre critère d'a priori géométrique. Certaines méthodes de segmentation par contours actifs [15, 16, 55, 69, 75] font alors l'hypothèse d'une transformation paramétrique entre le contour actif et le contour de référence. L'estimation des paramètres de la transformation est couplée au problème de segmentation. Une extension de nos travaux consisterait à définir un critère minimisant la distance entre le contour actif et la transformée, par rapport à un modèle paramétrique, du contour de référence. Dans l'esprit de la définition du critère de segmentation et d'estimation conjointe du mouvement, proposée dans ce document, le problème de minimisation de l'énergie serait conjoint et non couplé. Le problème couplé consiste à voir l'énergie comme fonction du contour et des paramètres de mouvement. On obtient itérativement une segmentation, en dérivant l'énergie par rapport au contour pour des paramètres fixés, et une estimation des paramètres, en minimisant l'énergie par rapport aux paramètres pour un contour fixé. Le problème conjoint consisterait à définir deux énergies dépendantes du contour : une pour la segmentation utilisant les paramètres de mouvement, et l'autre l'estimation de la transformation paramétrique. La minimisation de l'énergie de segmentation doit prendre en compte la dépendance des paramètres de la transformation au contour. En fonction du choix de l'énergie définissant les paramètres de la transformation, l'équation d'évolution du contour actif est loin d'être évidente. L'estimation des paramètres de la transformation permet également de recaler les images.

En outre, il serait intéressant d'étendre le critère d'a priori géométrique à la dimension supérieure. Le critère défini dans ce document minimise la distance du contour actif à un contour de référence. On pourrait minimiser la distance d'une surface active à une surface de référence. Il faudrait porter une attention particulière à la définition de distance entre surfaces introduite dans l'a priori géométrique. Son application à la segmentation et au suivi d'objets 3D étendrait celle proposée dans ce document pour la segmentation d'image et le suivi de régions dans une vidéo.

Cinquième partie

Annexe : Détection d'objets en mouvement

En l'absence de mouvement de caméra, les objets en mouvement représentent la source principale de différence entre deux images. Partant de ce constat simple, une segmentation courante de l'image consiste à définir :

- le fond de l'image comme la partie statique de l'image
- le ou les objets comme la partie mobile de l'image

Cette segmentation est donc une détection du mouvement.

Une des méthodes les plus simples de détection de mouvement consiste en l'étude de la différence entre deux images, la FD (*frame difference*), pour des séquences à caméra fixe. Afin d'étendre la segmentation à l'ensemble des vidéos, il faudra prendre en compte dans notre définition un éventuel mouvement de la caméra.

Pour les séquences à caméra mobile, on ne peut plus définir l'objet comme l'ensemble des pixels en mouvement, car le mouvement de la caméra induit un mouvement de tous les pixels de l'image. Cependant les pixels de l'objet ont un mouvement local propre qui s'ajoute à ce mouvement global. L'objet va donc être défini comme l'ensemble des pixels ne suivant pas le mouvement dominant de l'image.

Une gestion simple du mouvement de la caméra consiste à compenser le mouvement de la caméra, c'est-à-dire recaler l'une des images sur l'autre, pour s'approcher d'un cadre de caméra fixe. La détection s'appuie sur la DFD (*Displaced frame difference*), image différence des deux images, dont l'une est compensée en mouvement.

Toujours dans le but de séparer le fond de l'image des objets en mouvement, certaines méthodes s'appuient sur l'estimation d'une image représentant le fond statique de la scène, appelé fond de la séquence. Lorsque le fond de la séquence est connu, la détection des objets en mouvement est directe. Pour estimer le fond de la séquence, ces méthodes utilisent plusieurs images successives de la séquence.

Pour des séquences à caméra fixe, Jehan-Besson *et al.* [50] proposent un critère de détection de mouvement fondé sur la différence entre le fond estimé de la séquence considérée et l'image courante. Le fond d'une séquence de n images est considéré comme une moyenne de ces n images. Afin d'éliminer les valeurs correspondant aux objets en mouvement, les *outliers*, les auteurs introduisent dans leur critère un M-estimateur et minimisent la différence pondérée entre le fond et chacune des images par un algorithme de minimisations alternées.

Pour des séquences à caméra mobile, le fond de la séquence est appelé mosaïque. De nombreuses méthodes de construction de mosaïques sont présentées dans [44].

Nous proposons un descripteur utilisant le fond de la séquence ou mosaïque, afin d'étendre l'approche à caméra fixe, présentée par S. Jehan-Besson *et al.* dans [50], à des séquences à caméra mobile. Notre méthode repose sur deux étapes : l'estimation du fond de la séquence appelé mosaïque, puis la segmentation à partir du fond de l'image, extrait de la mosaïque.

Construire une mosaïque revient, pour chaque image, à extraire, même grossièrement, leur fond et à positionner ce fond par rapport à une image de référence. Cela suppose d'avoir une certaine idée de ce que sera le fond de chaque image, ce qui nécessite une segmentation préalable. Cette pré-segmentation n'a pas besoin d'être très précise : on la dira grossière, car on ne cherche pas à segmenter l'objet, mais juste à s'assurer que tous les pixels qui serviront à la construction de la mosaïque appartiennent au fond de chaque image. Pour cela, la pré-segmentation sera en fait une sur-segmentation fondée sur un descripteur classique s'appuyant sur deux images successives.

Ensuite, pour insérer dans la mosaïque les pixels du fond de chaque image, il faut compenser chacun des fonds, c'est-à-dire les placer dans le repère de l'image choisie comme image de référence. Compenser le fond d'une image par rapport à celle de référence nécessite de connaître le mouvement entre ces images.

Pour résumer, la construction de la mosaïque requiert les étapes suivantes :

- estimer et compenser le mouvement de la caméra entre deux images successives, l'image courante et sa précédente,
- pré-segmenter en utilisant un descripteur s'appuyant sur deux images : l'image courante et l'image précédente compensée en mouvement,
- estimer et compenser le mouvement de la caméra entre l'image courante et l'image de référence,
- insérer dans la mosaïque le fond de l'image courante par rapport à l'image de référence.

Une fois la mosaïque construite, nous disposons du fond de la séquence et du mouvement de la caméra entre chaque image et l'image de référence. Nous pouvons extraire de la mosaïque le fond de chaque image. Nous avons à présent tous les éléments nécessaires à la segmentation finale, fondée principalement sur la différence entre l'image et son fond.

L'ensemble des étapes-clés de notre méthode est résumé et illustré par la figure 8.

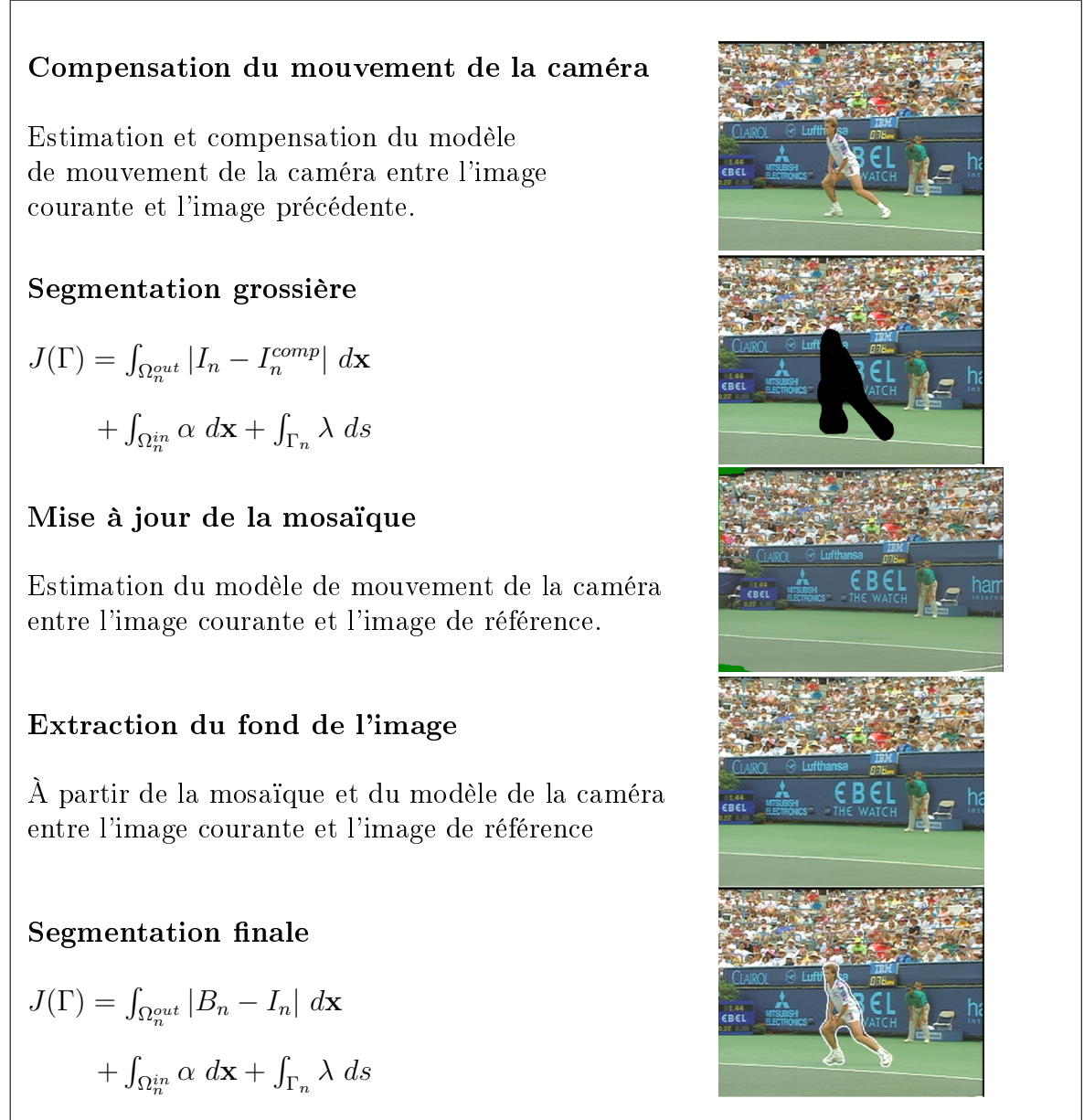


FIG. 8 – Étapes-clés de l'algorithme de détection de mouvement à partir d'une mosaïque.

Estimation du fond de la séquence : la mosaïque

La première étape de notre algorithme consiste à construire la mosaïque de la séquence choisie. L'intérêt pour les mosaïques va croissant et dépasse le simple effet visuel panoramique pour servir de base à une représentation complète et efficace de la séquence en compression vidéo ou en indexation.

Dans [44] les auteurs distinguent deux types de mosaïques : statiques et dynamiques. Les mosaïques statiques rendent compte du contenu statique de la séquence, c'est-à-dire exploitent les redondances temporelles. Les objets en mouvement vont de ce fait disparaître complètement, ou persister sous forme de traces résiduelles ou fantômes. Les différences entre l'image courante et la partie de la mosaïque qui lui correspond, appartiennent en majeure partie aux objets en mouvement.

Les mosaïques dynamiques, quant à elles, sont une suite de mosaïques évoluant à chaque nouvelle image, liées fortement à l'image en cours. Contrairement à la mosaïque statique, qui privilégie les informations redondantes (typiquement le fond des images), la mosaïque dynamique privilégie les données les plus récentes acquises à partir de l'image courante. Les différences entre l'image courante et la partie de la mosaïque qui lui est associée résulte donc de l'intégration des données des images précédentes.

Étant donné que nous souhaitons disposer du fond de chacune des images, nous nous sommes naturellement tournés vers une représentation statique du fond de la séquence.

La construction de la mosaïque induit une première estimation grossière du fond de l'image. Le fond de l'image est extrait par une segmentation visant à minimiser la différence entre l'image courante et l'image précédente compensée en mouvement. Il est donc nécessaire d'estimer et de compenser

le mouvement de la caméra entre l'image courante et l'image précédente. Nous avons choisi de représenter le mouvement de la caméra par un modèle de mouvement affine. Ensuite, il reste à recalculer le fond estimé de l'image courante dans le repère de l'image de référence, et à mettre à jour la mosaïque. Là encore, il est nécessaire d'estimer le modèle de mouvement de la caméra, cette fois entre l'image courante et l'image de référence.

Dans cette section, nous présenterons non pas les étapes successives de l'algorithme de construction de la mosaïque, mais les méthodes-clés mises en oeuvre. En effet, l'estimation et la compensation du mouvement de la caméra mettent en oeuvre les mêmes méthodes, qu'elles soient effectuées entre l'image courante et la précédente, ou entre l'image courante et celle de référence. Rappelons que les étapes de l'algorithme sont résumées dans la figure 8 page 133.

Les méthodes-clés qui seront détaillées par la suite concernent l'estimation du mouvement de la caméra (section A.1), la compensation du mouvement de la caméra (section A.2), l'extraction grossière du fond de chaque image (section A.3) et enfin la mise à jour de la mosaïque (section A.4).

A.1 Estimation du mouvement de la caméra

L'estimation du mouvement entre deux images repose principalement sur deux choix :

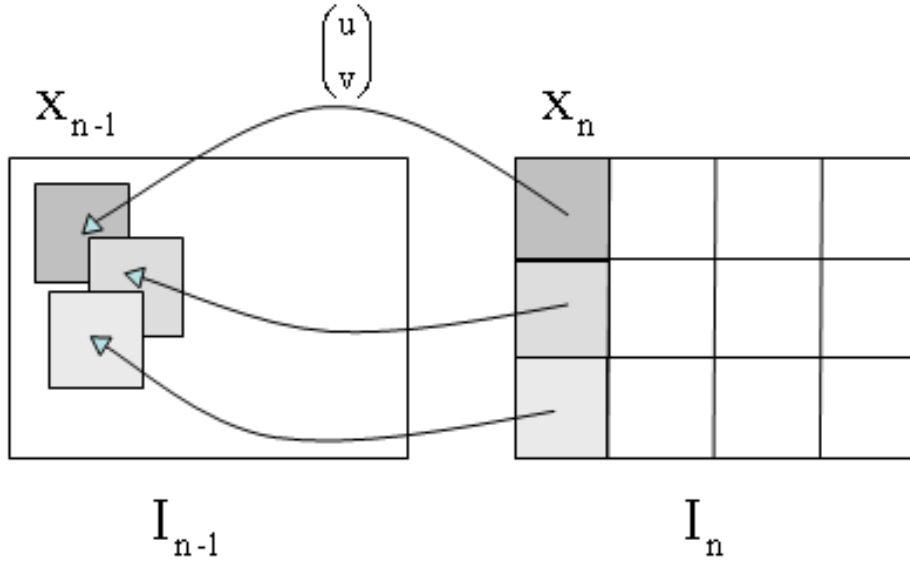
- le choix du modèle de représentation du mouvement,
- le choix de la méthode d'estimation du modèle de mouvement.

Les modèles de mouvement se classent en deux catégories :

- les modèles non paramétrés, qui reposent sur un champ de mouvement dense,
- les modèles paramétrés, qui représentent le mouvement d'une région par un jeu de paramètres.

Le mouvement de la caméra est, par hypothèse, le mouvement dominant de l'image, et les objets, par définition, sont les régions ayant un mouvement différent du mouvement dominant.

Puisqu'il s'agit d'estimer le mouvement de la région principale de l'image, les modèles paramétrés semblent donc tout indiqués. De plus, comme tous les pixels d'une région contribuent à l'estimation d'un jeu de paramètres, celle-ci gagne en robustesse et en précision.

FIG. A.1 – Estimation du mouvement dite *Backward*

Le modèle de mouvement choisi pour représenter le mouvement de la caméra est un modèle affine à 6 paramètres, car il représente un bon compromis entre la précision et le coût calculatoire [27].

La méthode d'estimation du modèle de mouvement se déroule en deux temps :

- l'image est découpée en blocs, et un algorithme de mise en correspondance ou *block-matching* associe un vecteur de mouvement à chaque bloc,
- un algorithme de minimisation définit le modèle de mouvement le plus représentatif de l'ensemble des vecteurs de mouvement.

A.1.1 Estimation du mouvement par mise en correspondance

Le principe du *block-matching* est le suivant : les images étant préalablement découpées en blocs, on cherche pour chaque bloc de l'une des deux images le bloc de l'autre image qui lui ressemble le plus. Le déplacement noté $\mathbf{v} = (u, v)^T$ du bloc se déduit de la différence d'emplacement des blocs considérés. Cette estimation peut se faire en recherchant les blocs correspondant à ceux de l'image courante, soit dans l'image suivante et la recherche est dite *forward*, soit dans l'image précédente et la recherche est dite *backward*.

Il existe plusieurs méthodes de recherche du bloc correspondant mais toutes s'effectuent dans une fenêtre de recherche définie par un paramètre de déplacement maximal.

Parmi les nombreuses méthodes de recherche de *block-matching*, citons le *block-matching* exhaustif, qui compare le bloc courant à tous les blocs contenus dans la fenêtre de recherche de l'autre image. Cette recherche exhaustive assure d'obtenir un minimum global pour la précision choisie, mais a pour principal inconvénient un coût calculatoire important.

Des méthodes de recherche alternatives tendent à réduire considérablement le coût calculatoire, tout en approchant au mieux la qualité d'estimation du *block-matching* exhaustif.

Citons l'algorithme précurseur de *three-step search* [53] qui recherche le bloc optimal dans un 8-voisinage à 4 pixels de distance, puis, autour du centre de ce bloc optimal, réitère la recherche pour un 8-voisinage à 2 pixels de distance et enfin, autour du nouveau bloc optimal, recherche le meilleur bloc dans un 8-voisinage à 1 pixel de distance.

Enfin Zhu *et al.* [92] proposent deux algorithmes itératifs améliorés dans le même esprit, utilisant des motifs de recherche non plus en carré, comme pour le *three-step search*, mais en diamant. Ces algorithmes améliorent la qualité de l'estimation par rapport au *three-step search* tout en réduisant le coût calculatoire.

Si le principal avantage des méthodes non exhaustives est leur coût calculatoire réduit, elles présentent toutefois le risque d'obtenir un minimum local.

Afin de pouvoir estimer les vecteurs de déplacement à une précision inférieure au pixel, l'image I_{n-1} a été interpolée par le filtrage H26L [47].

A.1.2 Estimation robuste du modèle de mouvement de la caméra

Représentons le modèle affine du mouvement de la caméra entre les images I_n et I_{n-1} par les matrices

$$A_{n,n-1} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \text{ et } \mathbf{t}_{n,n-1} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Notons \mathbf{x}_n un pixel de l'image I_n et \mathbf{x}_{n-1} son correspondant dans I_{n-1} au sens du *block-matching* tel que :

$$\mathbf{x}_{n-1} = \mathbf{x}_n + \mathbf{v}_{n,n-1} \quad (\text{A.1})$$

Le modèle de mouvement de la caméra doit représenter au mieux les différents déplacements de chacun des pixels de l'image I_n . Pour cela, nous

cherchons à minimiser la différence entre $A_{n,n-1}\mathbf{x}_n + \mathbf{t}_{n,n-1}$, l'image du pixel \mathbf{x}_n par le modèle de mouvement de la caméra, et \mathbf{x}_{n-1} son correspondant estimé par *block-matching*.

Un premier choix pour cette mesure de différence pourrait être la norme quadratique, car la solution du problème de minimisation est alors bien connue :

$$\sum_{\mathbf{x}_n \in I_n} (\mathbf{x}_{n-1} - A_{n,n-1}\mathbf{x}_n - \mathbf{t}_{n,n-1})^2 \quad (\text{A.2})$$

En effet, si l'on écrit la différence

$$\mathbf{x}_{n-1} - A_{n,n-1}\mathbf{x}_n - \mathbf{t}_{n,n-1} = \mathbf{x}_{n-1} - M_n \mathbf{p} \quad (\text{A.3})$$

avec

$$\mathbf{p} = (a_1 \ a_2 \ a_3 \ a_4 \ b_1 \ b_2)^T \quad (\text{A.4})$$

vecteur des paramètres de mouvement, et

$$M_n = \begin{pmatrix} x_n^i & x_n^j & 0 & 0 & 1 & 0 \\ 0 & 0 & x_n^i & x_n^j & 0 & 1 \end{pmatrix} \quad (\text{A.5})$$

avec $\mathbf{x}_n = (x_n^i, x_n^j)$, alors la solution du problème de la minimisation de (A.2) est donnée par :

$$\sum_{\mathbf{x}_n \in I_n} M_n^T M_n \mathbf{p} = \sum_{\mathbf{x}_n \in I_n} M_n^T \mathbf{x}_{n-1} \quad (\text{A.6})$$

Cependant un tel estimateur n'est pas robuste aux erreurs introduites par les objets en mouvement et les mises en correspondance aberrantes.

La mesure de différence φ choisie est un M-estimateur. Elle permet de pondérer l'influence des erreurs sur l'estimation du modèle de mouvement. Plus précisément, nous avons choisi la fonction de pénalisation de Geman et McClure [36] :

$$\varphi(r) = \frac{r^2}{1 + r^2}$$

Les courbes correspondantes $\varphi(r)$ et $\varphi'(r)/2r$ sont données Fig.A.2.

Le modèle de mouvement de la caméra est celui qui minimise le critère :

$$\sum_{\mathbf{x}_n \in I_n} \varphi(\mathbf{x}_{n-1} - A_{n,n-1}\mathbf{x}_n - \mathbf{t}_{n,n-1}) \quad (\text{A.7})$$

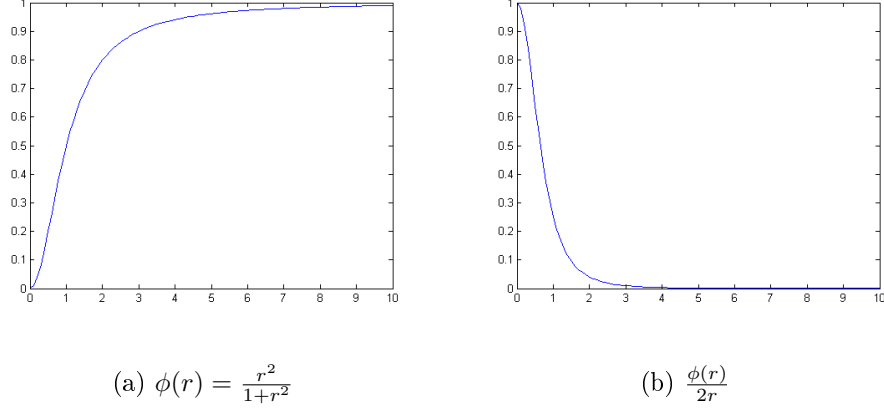


FIG. A.2 – L'estimateur de Geman et Mc Clure

Les M-estimateurs permettent de donner moins d'importance que la fonction carrée aux pixels pour lesquels la différence $|\mathbf{x}_{n-1} - A_{n,n-1}\mathbf{x}_n - \mathbf{t}_{n,n-1}|$ est élevée. Ces pixels présentent une différence élevée soit parce qu'ils appartiennent à un objet en mouvement et ne suivent donc pas le mouvement global de la caméra, soit parce que l'estimation par *block-matching* a été mauvaise. Une bonne estimation du modèle de mouvement de la caméra ne peut passer que par l'exclusion ou une prise en compte moindre de ces pixels dits *outlier*.

L'inconvénient d'une telle fonction de pondération est que le critère n'est plus différentiable directement. Pour minimiser le critère d'estimation du modèle de mouvement de la caméra nous avons utilisé le théorème énoncé par Charbonnier *et al.* [13] :

Théorème A.1. *Soit une fonction φ vérifiant les propriétés suivantes :*

1. φ est une fonction à valeur réelle continûment différentiable.
2. $\varphi(r) \geq 0 \quad \forall r$ avec $\varphi(0) = 0$.
3. φ est croissante sur \mathcal{R}^+ .
4. $\varphi(r) = \varphi(-r)$.
5. $\varphi'(r)/2r$ continue et strictement décroissante sur $[0, +\infty)$.
6. $\lim_{r \rightarrow +\infty} \frac{\varphi'(r)}{2r} = 0$.

$$7. \lim_{r \rightarrow 0^+} \frac{\varphi'(r)}{2r} = M, \quad \text{avec } 0 < M < +\infty.$$

Alors :

1. Il existe alors une fonction ψ strictement convexe et décroissante, $\psi : (0, M] \rightarrow [0, \beta)$ avec :

$$\beta = \lim_{r \rightarrow +\infty} \left(\varphi(r) - r^2 \frac{\varphi'(r)}{2r} \right)$$

et telle que :

$$\varphi(r) = \inf_{0 < w \leq M} (wr^2 + \psi(w))$$

2. Pour chaque r fixé, la valeur w_r pour laquelle le minimum est atteint, i.e., telle que :

$$\inf_{0 < w \leq M} (wr^2 + \psi(w)) = (w_r r^2 + \psi(w_r)),$$

est unique et donnée par :

$$w_r = \frac{\varphi'(r)}{2r}$$

Une démonstration de ce théorème est donnée dans [13, 3].

Remarquons que la fonction de Geman et Mc Clure vérifie les conditions du théorème A.1.

Le théorème A.1 nous permet d'affirmer que minimiser le critère (A.7) est équivalent à minimiser le critère semi-quadratique

$$\sum_{\mathbf{x}_n \in I_n} (b_{min}(\mathbf{x}_{n-1} - A_{n,n-1}\mathbf{x}_n - \mathbf{t}_{n,n-1})^2 + \psi(b_{min})) \quad (\text{A.8})$$

avec

$$b_{min} = \frac{\varphi'(|\mathbf{x}_{n-1} - A_{n,n-1}\mathbf{x}_n - \mathbf{t}_{n,n-1}|)}{2 * |\mathbf{x}_{n-1} - A_{n,n-1}\mathbf{x}_n - \mathbf{t}_{n,n-1}|} \quad (\text{A.9})$$

L'algorithme de minimisation consiste à calculer itérativement :

- les paramètres du modèle $A_{n,n-1}$ et $\mathbf{t}_{n,n-1}$ en minimisant l'équation (A.8), à b_{min} fixé
- les pondérations b_{min} à partir de (A.9)

La méthode d'estimation du mouvement de la caméra a été développée entre les images I_n et I_{n-1} , mais elle s'applique de même entre les images I_n et I_{ref} .

A.2 Compensation du mouvement de la caméra

La compensation du mouvement de la caméra entre les images I_n et I_{n-1} peut s'effectuer suivant deux stratégies opposées : *forward* et *backward*.

Le principe d'une compensation *forward* consiste, pour chaque pixel \mathbf{x}_{n-1} de l'image I_{n-1} à calculer son emplacement \mathbf{x}_{comp} dans l'image compensée I_{n-1}^{comp} , soit :

$$\mathbf{x}_{comp} = A_{n-1,n}\mathbf{x}_{n-1} + \mathbf{t}_{n-1,n} \quad (\text{A.10})$$

où $A_{n-1,n}$ et $\mathbf{t}_{n-1,n}$ représentent le modèle de mouvement de la caméra entre les images I_{n-1} et I_n .

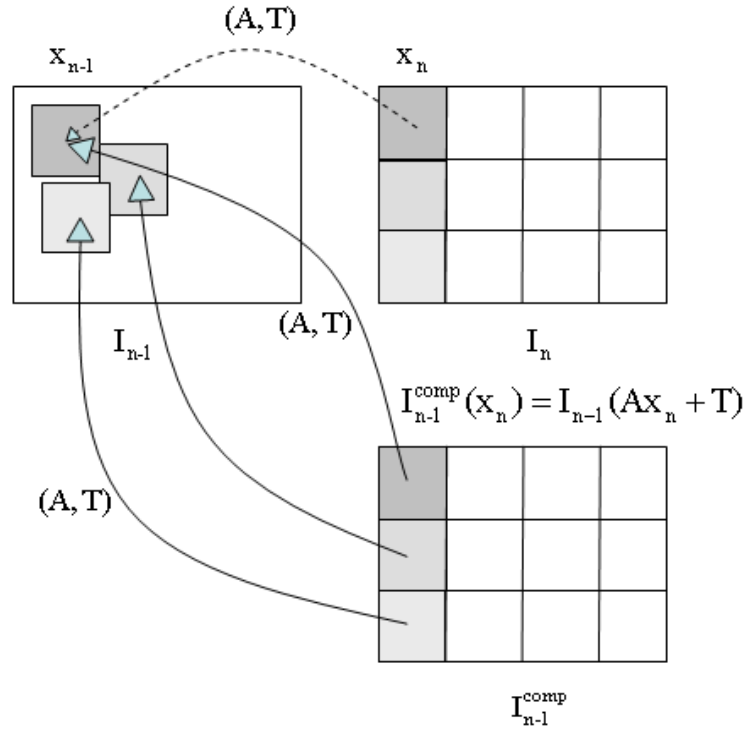


FIG. A.3 – Compensation *backward* du mouvement de la caméra entre les images I_n et I_{n-1}

Le pixel \mathbf{x}_{comp} ainsi déterminé se verra affecter l'intensité du pixel \mathbf{x}_{n-1} qui

lui correspond :

$$I_{n-1}^{comp}(\mathbf{x}_{comp} = A_{n-1,n}\mathbf{x}_{n-1} + \mathbf{t}_{n-1,n}) = I_{n-1}(\mathbf{x}_{n-1}) \quad (\text{A.11})$$

Cette méthode n'assure pas d'avoir une valeur pour chaque pixel de l'image compensée, qui risque donc de présenter des trous. De plus, plusieurs pixels de l'image I_{n-1} peuvent avoir la même image dans I_{n-1}^{comp} , créant des zones de recouvrement.

Le principe de la méthode *backward* consiste à assigner à chaque pixel \mathbf{x}_{comp} de l'image I_{n-1}^{comp} l'intensité du pixel \mathbf{x}_{n-1} qui lui correspond dans l'image I_{n-1} , pixel dont les coordonnées sont données par :

$$\mathbf{x}_{n-1} = A_{n,n-1}\mathbf{x}_{comp} + \mathbf{t}_{n,n-1} \quad (\text{A.12})$$

où $A_{n,n-1}$ et $\mathbf{t}_{n,n-1}$ représentent le modèle de mouvement de la caméra entre les images I_n et I_{n-1} .

On a donc :

$$I_{n-1}^{comp}(\mathbf{x}_{comp}) = I_{n-1}(\mathbf{x}_{n-1} = A_{n,n-1}\mathbf{x}_{comp} + \mathbf{t}_{n,n-1}) \quad (\text{A.13})$$

On s'assure ainsi que chaque pixel de l'image compensée se verra affecter une intensité. Deux pixels de l'image compensée peuvent toutefois avoir le même antécédent, créant des redondances d'information.

La figure A.3 présente une compensation de type *backward* telle que mise en oeuvre dans notre algorithme.

Afin d'augmenter la précision des résultats, les valeurs sub-pixelles des images ont été estimées à l'aide d'une interpolation H26L [47].

La méthode de compensation du mouvement de la caméra a été développée entre les images I_n et I_{n-1} mais elle s'applique de même entre les images I_n et I_{ref} .

A.3 Segmentation grossière du fond des images

Construire une mosaïque revient à recaler dans un même espace de référence différentes images représentant la même scène sous des points de vue différents. Si, d'une image à l'autre, des éléments de la scène ont été (ou se sont) déplacés, ces éléments se retrouveront sur la mosaïque sous forme de traces résiduelles. Afin d'éviter ces traces résiduelles, nous avons choisi de

pré-segmenter chaque image et de n'en garder que le fond.

La segmentation suit une approche variationnelle et est mise en oeuvre par un algorithme de contours actifs. Le critère de segmentation comprend trois termes :

$$\begin{aligned}
 J_n(\Gamma) &= \int_{\Omega_n^{out}} \sum_{i=Y,U,V} |I_n^i(\mathbf{x}) - I_{n-1}^{i,comp}(\mathbf{x})| d\mathbf{x} \\
 &+ \int_{\Omega_n^{in}} \alpha_c d\mathbf{x} \\
 &+ \int_{\Gamma_n} \lambda ds
 \end{aligned} \tag{A.14}$$

Ω_n^{out} et Ω_n^{in} représentent respectivement la région extérieure et intérieure au contour actif Γ_n de l'image I_n . Les variables α_c et λ sont deux constantes réelles.

À convergence, le contour actif détoure l'objet en mouvement, partitionnant l'image entre l'objet en mouvement défini par Ω_n^{in} et le fond de l'image défini par Ω_n^{out} .

Le premier terme du critère (A.14) évalue la différence entre le fond de l'image courante et celui de l'image précédente, compensée en mouvement. Si ce terme n'entrait en compétition avec aucun autre, le contour actif, évoluant de façon à minimiser le critère, tendrait à minimiser l'aire du fond de l'image, par-delà les contours de l'objet d'intérêt.

Le deuxième terme compense l'influence du premier en minimisant l'aire de l'objet.

Le troisième terme minimise la longueur du contour et permet d'avoir un contour plus régulier.

Les constantes α_c et λ pondèrent chacun des termes et permettent à l'utilisateur d'adapter l'algorithme aux caractéristiques des images.

La différentiation de chacun de ces termes est simple et nous permet d'obtenir l'équation d'évolution du contour actif :

$$\begin{cases} \frac{\partial \Gamma_n(p, \tau)}{\partial \tau} &= (\alpha_c - \sum_{i=Y,U,V} |I_n^i(\mathbf{x}) - I_n^{i,comp}(\mathbf{x})| + \lambda \cdot \kappa) \mathbf{N} \\ \Gamma_n(0) &= \Gamma_{n,0} \end{cases} \tag{A.15}$$

La courbure du contour est notée κ , $\mathbf{x} = \Gamma_n(p, \tau)$, le vecteur normal unitaire extérieur au contour est noté \mathbf{N} et le contour initial $\Gamma_{n,0}$ est un contour quelconque défini par l'utilisateur.



FIG. A.4 – Contours volontairement sur-segmentés des images 70, 75, 80 et 85.

L'équation d'évolution a été implémentée suivant la méthode des ensembles de niveaux [67].

La figure A.4 présente quelques contours obtenus en favorisant volontairement le premier terme du critère (A.14), ceci afin d'induire une sur-segmentation de l'objet et donc de s'assurer que le fond de l'image ne contient pas d'objet en mouvement.

A.4 Mise à jour de la mosaïque

Chacun des fonds segmentés à l'étape précédente doit être recalé dans un repère commun, défini par une image de référence I_{ref} . Nous avons choisi la première image de la vidéo comme image de référence.

Le recalage du fond d'une image passe par l'estimation et la compensation du mouvement de la caméra entre cette image et l'image de référence. Les modèles de mouvement calculés à cette étape sont précieusement conservés car ils serviront par la suite à extraire le fond de chaque image de la mosaïque (Section B.1).

Chaque fond apporte une quantité d'informations qui, intégrée à l'ensemble des informations déjà acquises, participe à la construction de la mosaïque. La plupart du temps, ces informations sont redondantes, ce qui nous permet d'ailleurs de les situer dans un repère commun. L'étape de mise à jour de la mosaïque détermine la façon dont ces informations vont s'intégrer à celles déjà présentes dans la mosaïque.

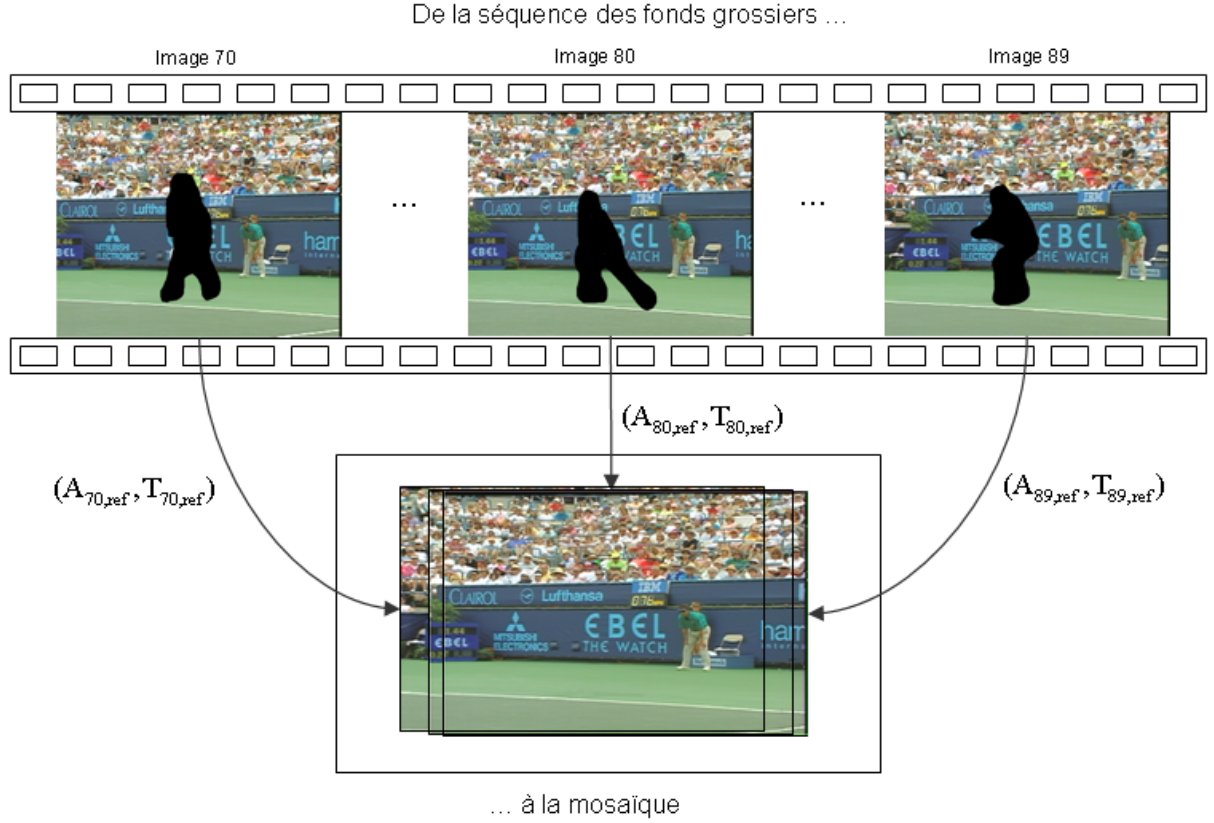


FIG. A.5 – Construction de la mosaïque à partir des fonds grossiers.

Afin de prendre en compte les différentes valeurs proposées par différentes

images pour un même pixel de la mosaïque, Nicolas *et al.* proposent dans [65] de mettre à jour la mosaïque à chaque nouvelle image :

$$M_n(\mathbf{x}) = (1 - \alpha) * M_{n-1}(\mathbf{x}) + \alpha I_n(\mathbf{x}) \quad (\text{A.16})$$

On note $M_n(\mathbf{x})$ la valeur de la mosaïque au pixel \mathbf{x} après intégration des données de l'image I_n . La constante α permet de pondérer l'influence de la nouvelle image sur la mosaïque.

Cependant, ce filtrage temporel est sensible à l'introduction de valeurs non représentatives. Nous avons utilisé une médiane temporelle des différentes valeurs attribuées à un pixel [44].

Segmentation finale : détection d'objets en mouvement avec fond estimé

Le fond de la séquence est connu sous la forme d'une mosaïque, il est aisé d'en déduire le fond de chaque image de la séquence. Nous développerons les deux dernières étapes de notre algorithme : l'extraction du fond de chaque image à partir de la mosaïque et la détection d'objets en mouvement.

B.1 Extraction du fond de chaque image à partir de la mosaïque

Le fond de chaque image se déduit de la mosaïque à l'aide des modèles de mouvement estimés à l'étape A.4, en utilisant une compensation *backward*.

Chaque pixel \mathbf{x}_n du fond B_n de l'image I_n a pour intensité celle du pixel \mathbf{x}_{ref} qui lui correspond dans la mosaïque M , par rapport au modèle de mouvement de la caméra entre l'image de référence et l'image courante :

$$B_n(\mathbf{x}_n) = M(\mathbf{x}_{ref} = A_{n,ref}\mathbf{x}_n + \mathbf{t}_{n,ref}) \quad (\text{B.1})$$

Ainsi nous estimons à partir de la mosaïque le fond correspondant à chaque image de la séquence.

B.2 Détection d'objet en mouvement

Enfin nous mettons en oeuvre un algorithme de segmentation par contours actifs dont le critère peut sembler similaire à celui de la pré-segmentation.

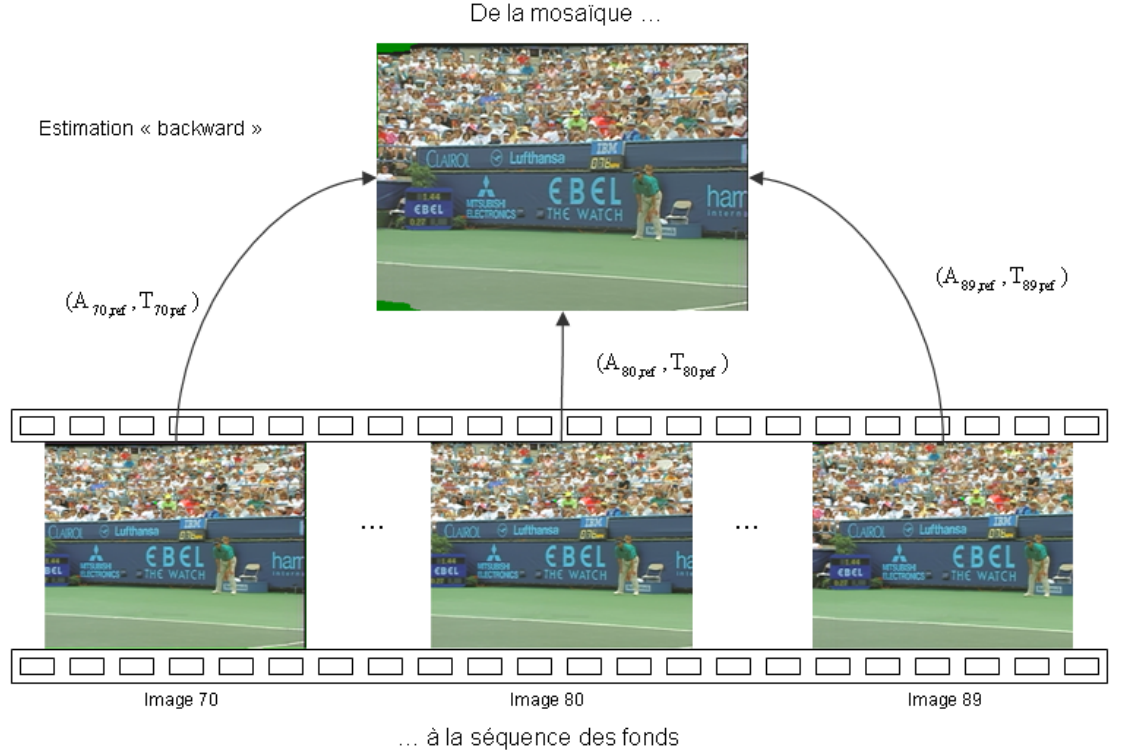


FIG. B.1 – Extraction des fonds de la séquence à partir de la mosaïque.

La différence se situe au niveau du terme décrivant la région extérieure au contour.

Ce descripteur, lors de la pré-segmentation, est fondé sur le gradient temporel entre deux images successives, dont l'une est compensée en mouvement. Pour la segmentation finale, ce descripteur repose sur la différence entre une image et le fond correspondant estimé à partir de plusieurs images de la séquence.

$$\begin{aligned}
 J_n(\Gamma) &= \int_{\Omega_n^{out}} \sum_{i=Y,U,V} |B_n^i(\mathbf{x}) - I_n^i(\mathbf{x})| d\mathbf{x} \\
 &+ \alpha_c \int_{\Omega_n^{in}} d\mathbf{x} + \lambda \int_{\Gamma_n} ds
 \end{aligned} \tag{B.2}$$

L'utilisation de plusieurs images dans le descripteur de la région extérieure au contour permet de rendre la segmentation plus robuste au bruit et

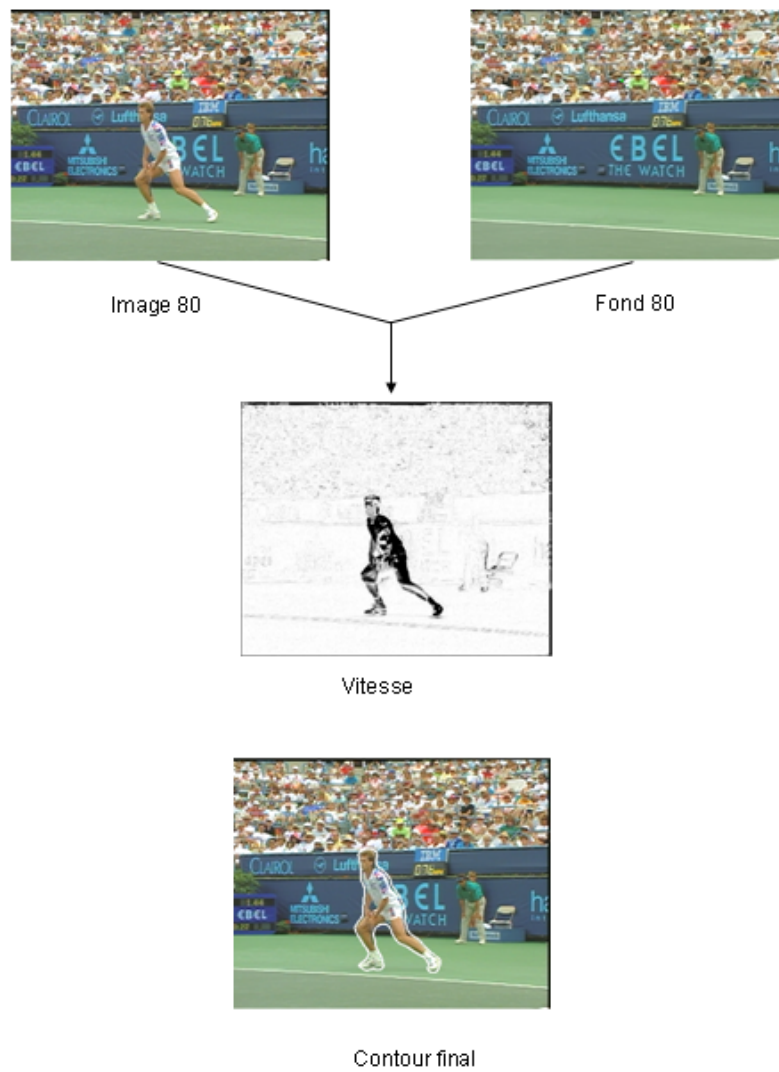


FIG. B.2 – Présentation du descripteur basé sur la différence entre l'image et son fond.

aux changements d'intensité. Ce descripteur permet également de ne pas considérer les zones d'occlusion comme faisant partie de l'objet, car le fond de l'image est connu.

L'équation d'évolution du contour actif est :

$$\begin{aligned}\frac{\partial \Gamma_n(\tau)}{\partial \tau} &= (\alpha_c - \sum_{i=1}^3 |B_n^i - I_n^i| + \lambda.\kappa) \mathbf{N} \\ \Gamma_n(0) &= \Gamma_{n,0}\end{aligned}\tag{B.3}$$

Elle a été implémentée par la méthode des ensembles de niveaux.

La figure B.2 présente une carte de vitesse du contour actif pour le critère B.2. On peut remarquer que la différence entre l'image et son fond laisse apparaître clairement l'objet en mouvement, ici le joueur de tennis. La dernière image présente le contour actif à convergence.

Annexe C

Résultats expérimentaux

Nous présentons les résultats obtenus pour la séquence **Stefan**. Cette séquence est particulièrement délicate à segmenter car les mouvements de la caméra incluent zooms et larges translations, et que le joueur de tennis se déplace rapidement.

La figure C.1 montre la mosaïque obtenue à partir de 20 images entre les images 70 et 89. L'objet se déplaçant rapidement, 20 images suffisent à construire une mosaïque complète, dans laquelle chaque zone recouverte par l'objet dans une image est découverte dans une autre image.

La figure C.2 présente, pour les images 70, 75, 77, 80 et 87, le fond de chaque image extrait à partir de la mosaïque et l'objet en mouvement détecté.

Dans la figure C.3, nous présentons les contours des images paires de l'extrait de 20 images de la séquence.

Afin d'effectuer une évaluation objective de la qualité de la segmentation, nous avons comparé les objets A^{est} , extraits par notre algorithme, à ceux de référence obtenus par segmentation manuelle.

Nous avons utilisé le critère de qualité spatiale de segmentation de Mech *et al.* [60] :

$$d_{abs} = \frac{\sum_{\mathbf{x} \in A^{est}} A^{ref}(\mathbf{x}) \oplus A^{est}(\mathbf{x})}{N} \quad (C.1)$$

où N est le nombre d'images de la séquence étudiée, et \oplus l'opérateur logique binaire du "ou exclusif".

La distorsion moyenne, d_{abs} , sur les 20 images de la séquence **Stefan** vaut 0.0090 ce qui signifie que moins de 1% des pixels sont mal étiquetés.



FIG. C.1 – Mosaïque estimée à partir de 20 images de la séquence Stefan (images 70 à 89)



FIG. C.2 – Fonds et objets des images 70,75,77, 80 et 87

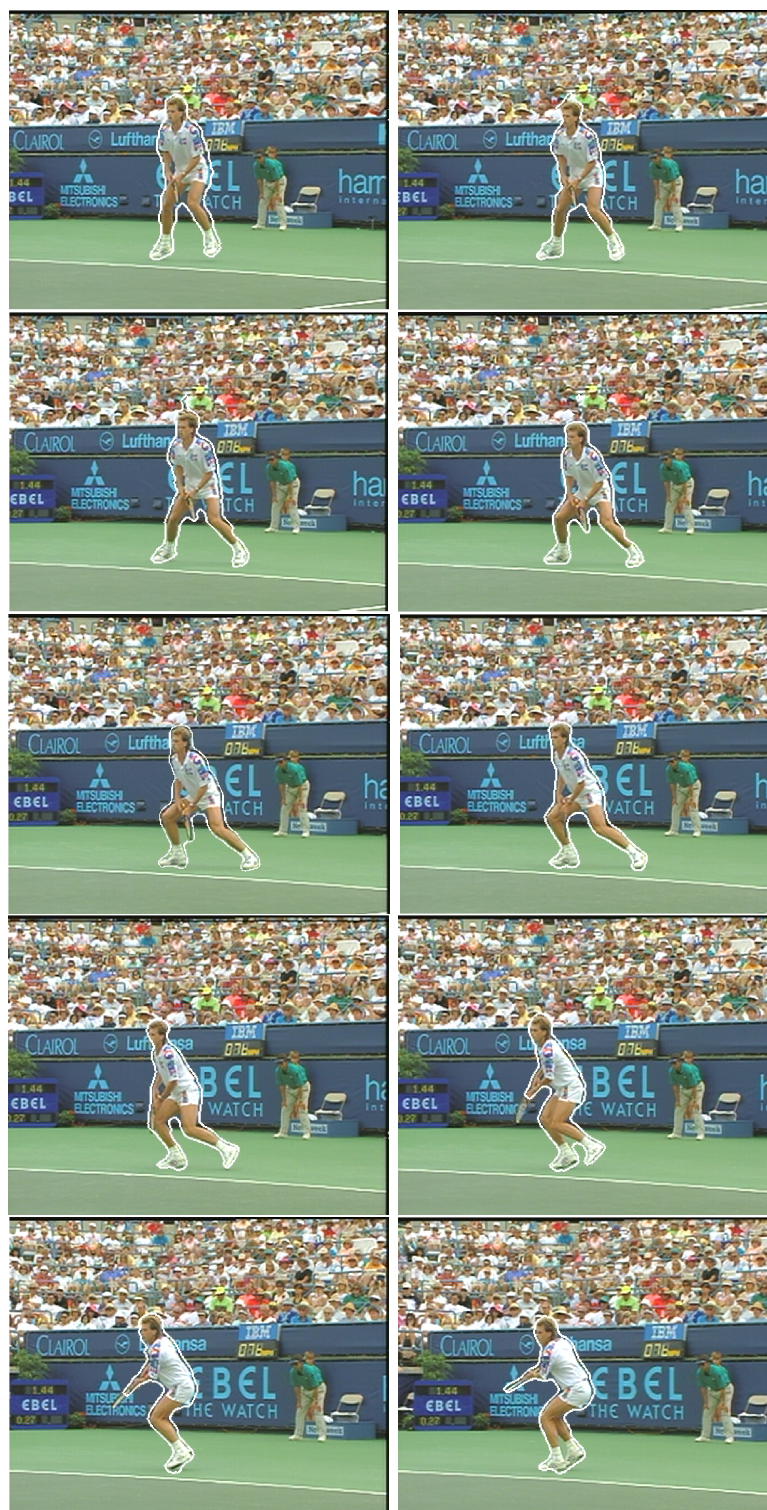


FIG. C.3 – Contours des images 70,72,74, 76, 78, 80, 82, 84, 86 et 88

Bibliographie

- [1] D. ADALSTEINSSON et J.A. SETHIAN. A fast level set method for propagating interfaces. *Journal of Computational Physics*, 118, 1995.
- [2] G. AUBERT, R. DERICHE, et P. KORNPORST. Computing optical flow via variational techniques. *SIAM Journal of Applied Mathematics*, 60(1) :156–182, 1999.
- [3] G. AUBERT et P. KORNPORST. *Mathematical problems in image processing. Partial differential equations and the calculus of variations*. Numéro 147. Applied Mathematical Sciences, Springer Verlag, 2001.
- [4] J.L. BARRON, D.J. FLEET, et S.S. BEACHEMIN. Performance of optical flow techniques. *International Journal on Computer Vision*, pages 43–77, 1994.
- [5] S. BEUCHER et F. MEYER. The morphological approach to segmentation : The watershed transformation. *Mathematical Morphology in Image Processing*, pages 433–481, 1993.
- [6] L. BLANC-FÉRAUD, M. BARLAUD, et T. GAIDON. Motion estimation involving discontinuities in a multiresolution scheme. *Optical Engineering*, 32(7) :1475–1482, 1993.
- [7] S. BOLTZ, É. DEBREUVE, et M. BARLAUD. A joint motion computation and segmentation algorithm for video coding. Dans *European Signal Processing Conference*, Antalya, Turquie, 2005.
- [8] X. BRESSON, P. VANDERGHEYNST, et J.P. THIRAN. A variational model for object segmentation using boundary information and statistical shape prior driven by the mumford-shah functional. *International Journal of Computer Vision*, accepté le 16 novembre 2005.
- [9] V. CASELLES, F. CATTE, T. COLL, et F. DIBOS. A geometric model for active contours in image processing. *Numerische Mathematik*, 66 :1–33, 1993.

- [10] V. CASELLES, R. KIMMEL, et G. SAPIRO. Geodesic active contours. *International Journal of Computer Vision*, 22(1) :61–79, 1997.
- [11] A. CHAKRABORTY, L. STAIB, et J. DUNCAN. Deformable boundary finding in medical images by integrating gradient and region information. *IEEE Transactions on Medical Imaging*, 15 :859–870, 1996.
- [12] T. CHAN et L. VESE. Active contours without edges. *IEEE Transactions on image processing*, 10(2) :266–276, 2001.
- [13] P. CHARBONNIER, L. BLANC-FÉRAUD, G. AUBERT, et M. BARLAUD. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2) :298–311, 1997.
- [14] G. CHARPIAT, O. FAUGERAS, et R. KERIVEN. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, pages OF1–OF58, 2004.
- [15] Y. CHEN, F. HUANG, H. TAGARE, M. RAO, D. WILSON, et E. GEISER. Using prior shape and intensity profile in medical image segmentation. Dans *IEEE International Conference on Computer Vision*, pages 1117–1125, Nice, France, 2003.
- [16] Y. CHEN, H. TAGARE, S. THIRUVENKADAM, F. HUANG, D. WILSON, K.S. GOPINATH, R.W. BRIGGS, et E.A. GEISER. Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3) :315–328, 2002.
- [17] L. COHEN. On active contour models and balloons. *Computer Vision, Graphics and Image Processing : Image Understanding*, 53 :211–218, 1991.
- [18] R. COURANT, K.O. FRIEDRICHS, et H. LEWY. On the partial difference equations of mathematical physics. *IBM Journal of Research and Development*, 11 :215–234, 1967.
- [19] D. CREMERS, C. SCHNORR, et J. WEICKERT. Diffusion-snakes : Combining statistical shape knowledge and image information in a variational framework. Dans *1st IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 137–144, Vancouver, Canada, 2001.
- [20] D. CREMERS et S. SOATTO. Variational space-time motion segmentation. Dans *International Conference on Computer Vision*, pages 886–893, 2003.
- [21] D. CREMERS et S. SOATTO. A variational framework for piecewise parametric motion segmentation. *International Journal of Computer Vision*, 62(3) :249–265, 2005.

- [22] D. CREMERS, F. TISCHHAÄUSER, J. WEICKERT, et C. SCHNÖRR. Diffusion-snakes : Introducing statistical shape knowledge into the mumford-shah functional. *International Journal of Computer Vision*, 7(6) :662–673, 2002.
- [23] É. DEBREUVE. *Segmentation par contours actifs en imagerie médicale dynamique : application en cardiologie nucléaire*. Thèse de doctorat, Université de Nice-Sophia Antipolis, France, 2000.
- [24] É. DEBREUVE, M. BARLAUD, G. AUBERT, Y. LAURETTE, et J. DARCOURT. Space time segmentation using level set active contours applied to myocardial gated SPECT. *IEEE Transactions on Medical Imaging*, 20(7) :643–659, 2001.
- [25] E. DEBREUVE, M. GASTAUD, M. BARLAUD, et G. AUBERT. A region-based joint motion computation and segmentation on a set of frames. Dans *6ème European Workshop on Image Analysis for Multimedia International Services*, Montreux, Suisse, avril 2005.
- [26] M.C. DELFOUR et J.-P. ZOLÉSIO. *Shapes and geometries : Analysis, differential calculus, and optimization*. Advances in Design and Control. Society for Industrial and Applied Mathematics (SIAM), 2001.
- [27] F. DUFAUX et F. MOSCHENI. *Video coding : a second generation approach*, chapitre Segmentation-based motion estimation for second generation video coding techniques. Kluwer Academic, Boston, 1996.
- [28] C. EPSTEIN et M. GAGE. *The curve shortening flow. Wave motion : theory, modelling, and computation*. Mathematical Sciences Research Institute Publications - Springer-Verlag, 1987.
- [29] R. FABLET, P. BOUTHEMY, et M. GELGON. Moving object detection in color image sequences using region-level graph labeling. Dans *IEEE International Conference on Image Processing*, Kobe, Japon, 1999.
- [30] A. FOULONNEAU, P. CHARBONNIER, et F. HEITZ. Affine-invariant geometric shape priors for region-based active contours. Rapport technique RR-AF01-2005, LRPS ERA 27/LSIIT UMR 7005, 2005. short version submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI).
- [31] F. GALLAND, N. BERTAUX, et P. RÉFRÉGIER. Multicomponent image segmentation in homogeneous regions by stochastic complexity minimization. *Pattern Recognition*, 38(11) :1926–1936, 2005.
- [32] V. GARCIA, É. DEBREUVE, et M. BARLAUD. A contour tracking algorithm for rotoscopy. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, Mai 2006.

- [33] M. GASTAUD et M. BARLAUD. Video segmentation using active contours on a group of pictures. Dans *IEEE International Conference on Image Processing*, volume 2, pages 81–84, Rochester, N.Y., septembre 2002.
- [34] M. GASTAUD, M. BARLAUD, et G. AUBERT. Tracking video objects using active contours and geometric priors. Dans *IEEE 4th European Workshop on Image Analysis for Multimedia International Services*, pages 170–175, Londres, RU, avril 2003.
- [35] M. GASTAUD, M. BARLAUD, et G. AUBERT. Combining shape prior and statistical features for active contour segmentation. *IEEE TCSVT special session on Audio and Video Analysis for Interactive Multimedia Services*, 14(5) :726–734, mai 2004.
- [36] S. GEMAN et D.E. MCCCLURE. Bayesian image analysis : an application to single photon emission tomography. Dans *Statistical Computing Section*, pages 12–18, Washington, DC, 1985.
- [37] A. HERBULOT, S. JEHAN-BESSON, S. DUFFNER, M. BARLAUD, et G. AUBERT. Segmentation of vectorial image features using shape gradients and information measures. *Soumis à Journal of Mathematical Imaging and Vision*, 2005.
- [38] M. HINTERMÜLLER et W. RING. A second order shape optimization approach for image segmentation. *SIAM, Journal of Applied Mathematics*, 64(2) :442–467, 2003.
- [39] B. K. P. HORN et B. G. SCHUNCK. Determining optical flow. *Artificial Intelligence*, 17 :185–203, 1981.
- [40] International Organisation for Standardization, ISO Standards : JTC 1/SC 29 – Coding of audio, picture, multimedia and hypermedia information. ISO/IEC 14496-2 :2004 : *Technologies de l'information – Codage des objets audiovisuels – Partie 2 : Codage visuel*, Ed.3, 2004.
- [41] International Organisation for Standardization, ISO Standards : JTC 1/SC 29 – Coding of audio, picture, multimedia and hypermedia information. ISO/IEC 15938-3 :2002 : *Technologies de l'information – Interface de description du contenu multimédia – Partie 3 : Visuel*, 2002.
- [42] International Organisation for Standardization, ISO Standards : JTC 1/SC 29 – Coding of audio, picture, multimedia and hypermedia information. ISO/IEC 11172-2 :1993 : *Technologies de l'information – Codage de l'image animée et du son associé pour les supports de stockage numérique jusqu'à environ 1,5 Mbit/s – Partie 2 : Vidéo*, 1993.
- [43] International Organisation for Standardization, ISO Standards : JTC 1/SC 29 – Coding of audio, picture, multimedia and hypermedia information. ISO/IEC 13818-2 :2000 : *Technologies de l'information –*

- Codage générique des images animées et du son associé : Données vidéo, Ed. 2, 2000.*
- [44] M. IRANI, P. ANANDAN, J. BERGEN, R. KUMAR, et S. HSU. Efficient representations of video sequences and their applications. *Signal Processing : Image Communication. Special issue on Image and Video Semantics : Processing, Analysis, and Application*, 8(4), 1996.
 - [45] ITU : International Telecommunication Union, Telecommunication Standardization Sector. *Video Codec For Audiovisual Services at p *64 kbit/s - ITU-T Recommendation H.261*, 1993.
 - [46] ITU : International Telecommunication Union, Telecommunication Standardization Sector. *Video Coding for low bitrate communication - ITU-T Recommendation H.263*, 1996.
 - [47] ITU : International Telecommunication Union, Telecommunications Standardization Sector, Study Group 16, Video Coding Experts Group, Austin, Texas USA. *H.26L Test Model Long Term Number 87*, 2001.
 - [48] S. JEHAN-BESSON. *Modèles de contours actifs basés régions pour la segmentation d'images et de vidéos*. Thèse de doctorat, Université de Nice-Sophia Antipolis, 2003.
 - [49] S. JEHAN-BESSON, M. BARLAUD, et G. AUBERT. Video object segmentation using eulerian region-based active contours. Dans *International Conference on Computer Vision*, pages 353–361, Vancouver, Canada, 2001.
 - [50] S. JEHAN-BESSON, M. BARLAUD, et G. AUBERT. A 3-step algorithm using region-based active contours for video objects detection. *EURASIP journal on applied signal processing*, 2002(6) :572–581, 2002.
 - [51] S. JEHAN-BESSON, M. BARLAUD, et G. AUBERT. DREAM² S : Deformable regions driven by an eulerian accurate minimization method for image and video segmentation. *International Journal of Computer Vision*, 53(1) :45–70, 2003.
 - [52] M. KASS, A. WITKIN, et D. TERZOPOULOS. Snakes : Active contour models. *International Journal of Computer Vision*, 1 :321–332, 1988.
 - [53] T. KOGA, K. LINUMA, A. HIRANO, Y. IJIMA, et T. ISHIGURO. Motion-compensated interframe coding for video conferencing. Dans *National Telecommunication Conference*, page G5.3.15, New Orleans, LA, 1981.
 - [54] P. KORNPORST, R. DERICHE, et G. AUBERT. Image sequence analysis via partial differential equations. *Journal of Mathematical Imaging and Vision*, 11(1) :5–26, 1999.

- [55] M. LEVENTON, E. GRIMSON, et O. FAUGERAS. Statistical shape influence in geodesic active contour. Dans *Computer Vision and Pattern Recognition*, pages 1316–1323, Hilton Head Island, South Carolina, 2000.
- [56] F. LEYMARIE et D. LEVINE. Tracking deformable objects in the plane using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6) :617–634, 1993.
- [57] R. LI, B. ZENG, et M.L. LIOU. A new threestep search algorithm for block estimation. *IEEE Transactions on Circuits Systems and Video Technology*, 4 :438–442, 1994.
- [58] R. MALLADI, J.A. SETHIAN, et B.C. VEMURI. Shape modeling with front propagation : a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17 :158–175, february 1995.
- [59] M. MAZIÈRE et F. CHASSAING. Segmentation and tracking of video objects : suited to content-based video indexing, interactive television and production systems. Dans *International Conference on Image Processing*, Vancouver, Canada, 2000.
- [60] R. MECH. *Description of COST 211 analysis model*. COST 211^{quat} simulation group, Dublin, 1998.
- [61] A. MITICHE, R. FEGHALI, et A. MANSOURI. Motion tracking as spatio-temporal motion boundary detection. *Robotics and Autonomous systems*, 43 :39–50, 2003.
- [62] E. MÉMIN et P. PÉREZ. Hierarchical estimation and segmentation of dense motion fields. *International Journal of Computer Vision*, 46(2) :129–155, 2002.
- [63] D. MUMFORD et J. SHAH. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42 :577–684, 1989.
- [64] H. NAGEL et W. ENKELMANN. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(5) :565 – 593, 1986.
- [65] H. NICOLAS. Optimal criterion for dynamic mosaicking. Dans *International Conference on Image Processing*, Kobe, Japan, 1999.
- [66] J.-M. ODOBEZ et P. BOUTHEMY. Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6(4) :348–365, 1995.
- [67] S. OSHER et J.A. SETHIAN. Fronts propagating with curvature-dependent speed : Algorithms based on hamilton-jacobi formulation. *Journal of Computational Physics*, 79 :12–49, 1988.

- [68] C. PAPIN, P. BOUTHEMY, E. MÉMIN, et G. ROCHARD. Tracking and characterization of highly deformable cloud structures. Dans *European Conference on Computer Vision*, volume 2, pages 428–442, Dublin, Irlande, 2000.
- [69] N. PARAGIOS. Shape-based segmentation and tracking in cardiac image analysis. *IEEE Transactions on Medical Image Analysis*, pages 402–407, 2003.
- [70] N. PARAGIOS et R. DERICHE. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3) :266–280, 2000.
- [71] F. PRECIOSO. *Contours actifs paramétriques pour la segmentation d’images et vidéos*. Thèse de doctorat, Université de Nice-Sophia Antipolis, France, 2004.
- [72] F. PRECIOSO, M. BARLAUD, T. BLU, et M. UNSER. Robust real-time segmentation of images and videos using a smooth-spline snake-based algorithm. *IEEE Transactions on Image Processing*, 14(7) :910–924, 2005.
- [73] F. RANCHIN et F. DIBOS. Moving objects segmentation using optical flow estimation. Dans *Mathematics, Image and Analysis*, Paris, France, 2004.
- [74] R. RONFARD. Region-based strategies for active contour models. *International Journal of Computer Vision*, 13(2) :229–251, 1994.
- [75] M. ROUSSON et N. PARAGIOS. Shape priors for level set representations. Dans *European Conference in Computer Vision*, pages 78–92, Copenhagen, Danemark, 2002.
- [76] T. ROY, E. DEBREUVE, M. BARLAUD, et G. AUBERT. Segmentation of a vector field : Dominant parameter and shape optimization. *À paraître dans Journal of Mathematical Imaging and Vision*, 2006.
- [77] P. SALEMBIER et F. MARQUÉS. Region-based representations of image and video : Segmentation tools for multimedia services. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(8) :1147–1169, 1999.
- [78] C. SCHNÖRR. Computation of discontinuous optical flow by domain decomposition and shape optimization. *International Journal of Computer Vision*, 8(2) :153–165, 1992.
- [79] S. SOATTO et A. J. YEZZI. Deformation : Deforming motion, shape average and the joint registration and segmentation of images. Dans *European Conference on Computer Vision*, pages 32–47, Copenhagen, Danemark, 2002.

- [80] J. SOKOLOWSKI et J.-P. ZOLÉSIO. *Introduction to shape optimization. Shape sensitivity analysis.*, volume 16 de *Springer Ser. Comput. Math.* Springer-Verlag, Berlin, 1992.
- [81] L.H STAIB et J.S. DUNCAN. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 14(11) :1061–1075, 1992.
- [82] M. SUSSMAN, P. SMEREKA, et S. OSHER. A level set method for computing solutions of incompressible two-phase flows. *Journal of Computational Physics*, 114 :146–159, 1994.
- [83] A. TSAI, A. YEZZI, W. WELLS, C. TEMPANY, D. TUCKER, A. FAN, W.E. GRIMSON, et A. WILLSKY. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2) :137–154, 2003.
- [84] C. VACHIER et F. MEYER. The viscous watershed transform. *Journal of Mathematical Imaging and Vision*, 22(2-3) :251–267, 2005.
- [85] T. VEIT, F. CAO, et P. BOUTHEMY. A maximality principle applied to a *contrario* motion detection. Dans *IEEE International Conference on Image Processing*, Gènes, Italie, 2005.
- [86] P. VILLEGAS, X. MARICHAL, et A. SALCEDO. Objective evaluation of segmentation masks in video sequences. Dans *Workshop on Image Analysis for Multimedia International Services*, pages 85–88, Berlin, Allemagne, 1999.
- [87] Y. WANG et L. STAIB. Boundary finding with correspondence using statistical shape models. Dans *IEEE Conference in Computer Vision and Pattern Recognition*, pages 338–345, Santa Barbara, C.A., 1998.
- [88] J. WEICKERT et C. SCHNÖRR. Variational optic flow computation with a spatio-temporal smoothness constraint. *Journal of Mathematical Imaging and Vision*, 14(3) :245–255, 2001.
- [89] S. F. WU et J. KITTLER. A gradient-based method for general motion estimation and segmentation. *Journal of Visual Communication and Image Representation*, 4(1) :25–38, 1993.
- [90] L. Q. XU, J. L. LANDABASO, et M. PARDÀS. Shadow removal with blob-based morphological reconstruction for error correction. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 235–246, Pennsylvania , USA, 2005.
- [91] H-K ZHAO, S. OSHER, et R. FEDKIW. Fast surface reconstruction using the level set method. Dans *IEEE workshop on Variational and Level Set Methods in computer vision*, pages 194–203, 2001.

- [92] S. ZHU et K.-K. MA. A new diamond search algorithm for fast block matching motion estimation. *IEEE Transactions on Image Processing*, 9(2) :287–290, 2000.
- [93] S. ZHU et A. YUILLE. Region competition : unifying snakes, region growing, and bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18 :884–900, 1996.

Publications

Revue internationale avec comité de lecture

M. GASTAUD, M. BARLAUD, et G. AUBERT. Combining shape prior and statistical features for active contour segmentation. *IEEE TCSVT special issue on Audio and Video Analysis for Interactive Multimedia Services*, volume 14, numéro 5, pages 726–734, mai 2004.

S. JEHAN-BESSON, M. GASTAUD, F. PRECIOSO, M. BARLAUD, G. AUBERT, et É. DEBREUVE. From snakes to region-based active contours defined by region-dependent parameters. *Applied Optics*, volume 43, numéro 2, pages 247–256, janvier 2004.

Conférences internationales avec comité de lecture

É. DEBREUVE, M. GASTAUD, M. BARLAUD, et G. AUBERT. A region-based joint motion computation and segmentation on a set of frames. Dans *6th European Workshop on Image Analysis for Multimedia International Services*, Montreux, Suisse, avril 2005.

T. ANDRÉ, B. PESQUET-POPESCU, M. GASTAUD, M. ANTONINI, et M. BARLAUD. Motion estimation using chrominance for wavelet-based video coding. Dans *Proc. IEEE Picture Coding Symposium*, San Francisco, USA, décembre 2004.

S. JEHAN-BESSON, M. GASTAUD, M. BARLAUD, et G. AUBERT. Region-based active contours using geometrical and statistical features for image segmentation. Dans *IEEE International Conference on Image Processing*, vol. 2, Barcelone, Espagne, septembre 2003, pages 643–646.

M. GASTAUD, M. BARLAUD, et G. AUBERT. Tracking video objects using

active contours and geometric priors. Dans *IEEE 4th European Workshop on Image Analysis for Multimedia International Services*, avril 2003, pages 170–175.

M. GASTAUD, M. BARLAUD, et G. AUBERT. Tracking video objects using active contours. Dans *IEEE Workshop on Motion and Video Computing*, Orlando, Floride, USA, décembre 2002, pages 90–95.

M. GASTAUD et M. BARLAUD. Video segmentation using active contours on a group of pictures. Dans *IEEE International Conference on Image Processing*, volume 2, Rochester, New York, USA, septembre 2002, pages 81–84.

Rapports de recherche

M. GASTAUD, M. BARLAUD, et G. AUBERT. Tracking video objects using active contours and geometric priors. Rapport technique, CReATiVe, I3S/RR-2003-07-FR avril 2003.

M. GASTAUD, M. BARLAUD, et G. AUBERT. Combining geometric prior and statistical features for active contour segmentation. Rapport technique, CReATiVe, I3S/RR-2003-10-FR, mai 2003.

RÉSUMÉ

La segmentation en objets d'une image consiste à extraire de l'image des régions d'intérêt suivant un critère défini. Nous segmentons l'image par un algorithme de contours actifs dans le cadre d'une approche variationnelle. Partant d'un contour initial quelconque, le contour actif évolue, suivant une équation aux dérivées partielles. L'équation d'évolution du contour actif est déduite de la dérivation du critère. Au vu de la dépendance du critère à la région considérée, la dérivation du critère par rapport à la région n'est pas aisée. Nous utilisons des outils de dérivation empruntés à l'optimisation de domaine : les gradients de forme.

La contribution de cette thèse réside dans l'élaboration et l'étude de différents descripteurs de région. Pour chaque critère, nous calculons la dérivée du critère à l'aide des gradients de forme, et en déduisons l'équation d'évolution du contour actif. Le premier descripteur définit un a priori géométrique sans contrainte paramétrique : il minimise la distance du contour actif à un contour de référence. Nous l'avons appliqué à la déformation de courbe, la segmentation et le suivi de cible. Le deuxième descripteur caractérise le mouvement de l'objet par un modèle de mouvement. Le critère associé définit conjointement une région et son mouvement sur plusieurs images consécutives. Nous avons appliqué ce critère à l'estimation et la segmentation conjointe du mouvement et au suivi d'objets en mouvement.

Mots clés : Segmentation, contours actifs, minimisation, méthode variationnelle, équation aux dérivées partielles, gradients de forme, a priori, contour de référence, estimation et segmentation conjointe du mouvement, déformation de contours, suivi.

ABSTRACT

Image segmentation is the partitioning of an image into regions of interest and a background. Image segmentation can be performed using an active contour algorithm in a variational framework. In the case of a single region of interest, the active contour evolves from an initial contour towards the region of interest according to an evolution equation. In the variational framework, the evolution equation is deduced from the derivative of a criterion. The criterion should characterize the region of interest in terms of, for example, its motion, its color homogeneity or its texture. We differentiate the criterion using shape gradients, a theory originally developed for shape optimisation. We propose two criteria and we provide for each criterion the development leading to the associated evolution equation. The first criterion defines a free form a priori constraint on the contour by minimizing the distance between the active contour and a reference contour. We applied this criterion to shape warping, image segmentation, and video tracking. The second criterion assumes that the object motion can be described by a motion model on a group of pictures. The evolution of the active contour allowing to minimize this criterion provides a joint motion segmentation and motion estimation. We applied this criterion to sequence segmentation and to video tracking.

Keywords : Segmentation, active contours, minimization, variational framework, shape gradients, regularization, contour of reference, joint motion segmentation and estimation, shape warping, tracking.